

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds
(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524d-30

Perfect score: 79
Sequence: 1 MNIKLKMPTYIAGYK 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_73: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	374	2 JC2124	major allergen Cry
2	79	100.0	374	2 JC2123	major allergen Cry
3	44	55.7	1202	2 T23429	hypothetical prote
4	43	54.4	345	2 T28026	hypothetical prote
5	41	51.9	252	2 H26965	short-chain dehydr
6	40	50.6	522	2 H86248	protein T23J18.22
7	39	49.4	67	2 S20967	hypothetical prote
8	39	49.4	162	2 B18177	hypothetical pili
9	39	49.4	294	2 H38888	COI intron 15 prot
10	39	49.4	347	2 S59837	probable membrane
11	39	49.4	395	2 AH0441	probable exported
12	39	49.4	528	2 JC1118	alkyl sulfatase (E
13	39	49.4	757	2 T16609	hypothetical prote
14	38	48.1	149	2 S48927	hypothetical prote
15	38	48.1	268	2 A83651	hypothetical prote
16	38	48.1	320	2 P86821	hypothetical prote
17	38	48.1	541	2 DB2302	hypothetical prote
18	38	48.1	614	2 T16303	iron(III) ABC tran
19	38	48.1	646	1 T26038	probable membrane
20	38	48.1	647	2 T2579	68.6K capsid prote
21	37	46.8	157	2 H87601	transcription elon
22	37	46.8	166	1 KRB02B	keratin, 68K type
23	37	46.8	167	2 S21359	keratin, type I, c
24	37	46.8	182	1 KRB02A	keratin, 68K type
25	37	46.8	182	2 A02946	keratin, 59K type
26	37	46.8	300	2 AC2235	hypothetical prote
27	37	46.8	338	2 G69413	conserved hypothe
28	37	46.8	348	2 C96673	gamma-tocopherol m
29	37	46.8	363	2 B97199	uncharacterized co

30	37	46.8	370	2 A25004	keratin, 53K type
31	37	46.8	384	2 I61769	keratin 6d, type I
32	37	46.8	390	2 A60093	cytochrome c
33	37	46.8	400	2 AD3293	opgC protein limpo
34	37	46.8	475	2 T45766	hypothetical prote
35	37	46.8	483	2 A34720	keratin 8, type II
36	37	46.8	493	2 H72165	A18R protein - var
37	37	46.8	493	2 A36850	hypothetical prote
38	37	46.8	493	2 T28560	keratin, type II c
39	37	46.8	502	2 A23547	keratin, type II c
40	37	46.8	511	2 A39340	neurofilament prot
41	37	46.8	513	2 S08381	keratin, 58K type
42	37	46.8	520	2 JS0291	intermediate filam
43	37	46.8	524	2 A23518	keratin, 57K type
44	37	46.8	553	2 I59009	epidermal keratin
45	37	46.8	555	2 T20137	hypothetical prote

ALIGNMENTS

RESULT 1

JC2124
major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar
C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 14-Jul-1994 #sequence_reviseion 14-Jul-1994 #text_change 21-Jul-2000
C/Accession: JC2124

R/Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japan
A/Reference number: JC2123; PMID:94183234; PMID:8135802

A/Accession: JC2124

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-references: GB:D26545; NID:9493633; PIDN:BA05543.1; PID:9493634

A/Experimental source: pollen

A/Note: the authors described carbohydrate binding site for residue 279

C/Suprafamily: pectate lyase IAT59

C/Keywords: glycoprotein; pollen

F1-21/Domain: signal sequence #status predicted <SIG>

F/22-374/Product: major allergen Cry j I (clone pCCI-15) #status predicted <MAT>

F/158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 79; DB 2; Length 374;

Best local similarity 100.0%; Pred. No. 9.7e-07; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNIKLKMPTYIAGYK 15

DB 97 MNIKLKMPTYIAGYK 111

RESULT 2

JC2123
major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 14-Jul-1994 #sequence_reviseion 14-Jul-1994 #text_change 21-Jul-2000
C/Accession: JC2123; PMID:94183234; PMID:8135802

R/Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japan
A/Reference number: JC2123; PMID:94183234; PMID:8135802

A/Accession: JC2123

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-references: GB:D26544; NID:9493631; PIDN:BA05542.1; PID:9493632

A/Experimental source: pollen

A/Accession: PC2065

A/Molecule type: protein

A/Residues: 22-53;58-81;219-232;236-258;299-307;346-372 <SO2>

A/Note: the authors described carbohydrate binding site for residue 279

C/Suprafamily: pectate lyase IAT59

C/Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <STGS>
 F:22-37/Product: major allergen Cry I (clone PCC1-2-2) #status predicted <MAY>
 F:158,151,253,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 79; DB 2; Length 374;
 Best Local Similarity 100.0%; Pred. No. 9.7e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNIKKMPMYIAGYK 15
 DB 97 MNIKKMPMYIAGYK 111

RESULT 3

hypothetical protein K07G5.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C/Accession: T23429

R/McMurray, A.

submitted to the EMBL Data Library, April 1996

A/Reference number: Z19739

A/Accession: T23429

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1202 <WTL>

A/Cross-references: EMBL:Z71264; PIDN:CAA95830.1; GSPDB:GN00019; CESP:K07G5.3

A/Experimental source: clone K07G5

C/Genetics:

A/Map position: 1

A/Introns: 13/1; 104/3; 196/3; 260/2; 282/3; 307/2; 361/2; 410/2; 535/2; 603/2; 688/1; 7

C/Superfamily: Caenorhabditis elegans hypothetical protein K07G5.3

Query Match 55.7%; Score 44; DB 2; Length 1202;
 Best Local Similarity 53.8%; Pred. No. 11;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 IKLKMPMYIAGYK 15
 DB 811 IRLNIPLYISNYK 823

RESULT 4

hypothetical protein ZK829.8 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C/Accession: T28026

R/Harris, B.

submitted to the EMBL Data Library, May 1996

A/Reference number: Z20458

A/Accession: T28026

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-345 <MLT>

A/Cross-references: EMBL:Z73899; PIDN:CAA98078.1; GSPDB:GN00022; CESP:ZK829.8

A/Experimental source: clone ZK829

C/Genetics:

A/Map position: 4

A/Introns: 64/3; 120/2; 188/3; 277/2

Query Match 54.4%; Score 43; DB 2; Length 345;
 Best Local Similarity 53.8%; Pred. No. 4.3;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 IKLKMPMYIAGYK 15
 DB 30 LKFKSPYIIGYK 42

RESULT 5

H96965

short-chain dehydrogenase (gene dltE) [imported] - Clostridium acetobutylicum

C/Species: Clostridium acetobutylicum

C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C/Accession: H96965

R/Molling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Le

J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A/Reference number: A96900; MUID:21359325; PMID:21359325

A/Accession: H96965

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-252 <KUR>

A/Cross-references: GB:AE001437; PIDN:AAK78515.1; PID:G15023401; GSPDB:GN00168

A/Experimental source: Clostridium acetobutylicum ATCC824

C/Genetics:

A/Map position: 36

Query Match 51.9%; Score 41; DB 2; Length 252;
 Best Local Similarity 53.8%; Pred. No. 7.3;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNIKKMPMYIAG 13
 DB 109 VDNLKAPYIAG 121

RESULT 6

protein T23J18.22 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: H86248

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, T.

ansen, N.F.; Hughes, B.; Huizart, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Malt, R.; Marzita,

Rizzo, M.; Rooney, T.; Kowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: H86248

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-522 <STO>

A/Cross-references: GB:AE005172; NID:G6554189; PIDN:AAF16635.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

Query Match 50.6%; Score 40; DB 2; Length 522;
 Best Local Similarity 57.1%; Pred. No. 24;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 NIKLKMPMYIAGYK 15
 DB 240 NIREMPYIKDYE 253

RESULT 7

hypothetical protein 3 - Chlamydomonas reinhardtii chloroplast

C/Species: chloroplast Chlamydomonas reinhardtii

C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 18-Jun-1993

C/Accession: S20967

R/Croquet, Y.; Rahire, M.; Girard-Bascou, J.; Erickson, J.; Rochaix, J.D.

EMBO J. 11, 1697-1704, 1992

A/Title: A chloroplast gene is required for the light-independent accumulation of chl

A/Reference number: S20965; MUID:92258378; PMID:11374710

A:Accession: S20967
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-67 <CHO>
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 49.4%; Score 39; DB 2; Length 67;
Best Local Similarity 41.7%; Pred. No. 4.1;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 2 NIKMKPMYIAGY 13
DB 12 NIKMKPMYIAGY 23

RESULT 8
B81877
probable piliN NMA1110 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: B81877
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, N.; Holtz, S.; Jorgensen, K.; Leather, S.; Moutre, S.; Mungall, K.; Quail, M.A.; Rajandream, A.; Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: B81877
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162 <PAR>
A:Cross-references: GB:AL162755; GB:AL157959; NID:G7379742; PIDN:CA84372.1; PID:G737980
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1110

Query Match 49.4%; Score 39; DB 2; Length 162;
Best Local Similarity 46.2%; Pred. No. 11;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 3 IKMKPMYIAGY 15
DB 72 IKMKPMYIAGY 84

RESULT 9
H38888
COI intron 15 protein - Podospira anserina mitochondrion
C:Species: mitochondrion Podospira anserina
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: H38888
R:Camming, D.J.; Michel, F.; McNally, K.L.
Curr. Genet. 16, 381-406, 1989
A:Title: DNA sequence analysis of the 24.5 kilobase pair cytochrome oxidase subunit I mt
A:Reference number: A48327; MUID:90124722; PMID:2558609
A:Accession: H38888
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <CUM>
A:Cross-references: GB:X55026; GB:M30937; GB:M61734; NID:G14030; PIDN:CAA38793.1; PID:G1
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC3
C:Keywords: mitochondrion

Query Match 49.4%; Score 39; DB 2; Length 294;
Best Local Similarity 53.8%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 NIKMKPMYIAGY 14
DB 191 NIKMKPMYIAGY 203

RESULT 10

probable membrane protein YPR180W - Yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein P9705.5
C:Species: Saccharomyces cerevisiae
C>Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 06-Feb-1998
C:Accession: S59837
R:Paulley, A.
submitted to the EMBL Data Library, April 1995
A:Description: The sequence of S. cerevisiae cosmid 9705.
A:Reference number: S59829
A:Accession: S59837
A:Molecule type: DNA
A:Residues: 1-347 <PAU>
A:Cross-references: EMBL:U25842; NID:G786312; PID:G786317; MIPS:YPR180W
C:Genetics:
A:Gene: SGD:AOS1
A:Cross-references: SGD:S0006384; MIPS:YPR180W
A:Map position: 16R
C:Keywords: transmembrane protein
F:306-322/Domain: transmembrane #status predicted <TM>

Query Match 49.4%; Score 39; DB 2; Length 347;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 IKMKPMYIAGY 13
DB 145 IKMKPMYIAGY 154

RESULT 11

probable exported protein YPO3631 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AH0441
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.; deno-farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.; 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001, MUID:21470413; PMID:11586360
A:Accession: AH0441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-395 <KUR>
A:Cross-references: GB:AL590842; PIDN:CA093100.1; PID:G15981552; GSPDB:GN00175
C:Genetics:
A:Gene: YPO3631

Query Match 49.4%; Score 39; DB 2; Length 395;
Best Local Similarity 54.5%; Pred. No. 28;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 5 IKMKPMYIAGY 15
DB 320 IKMKPMYIAGY 330

RESULT 12

JC1118
alkyl sulfatase (EC 3.1.6.-) - Pseudomonas sp.
N:Alternate names: SDSase, sdsA protein
C:Species: Pseudomonas sp.
C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 08-Oct-1999
C:Accession: JC1118; S27644
R:Davidson, J.; Brunel, F.; Phanopoulos, A.; Prozzi, D.; Terpestra, P.
Gene 114, 19-24, 1992
A:Title: Cloning and sequencing of Pseudomonas genes determining sodium dodecyl sulfate
A:Reference number: JC1118; MUID:92267380; PMID:1587481

A:Accession: JCI118
 A:Molecule type: DNA
 A:Residues: 1-528 <DAV>
 A:Cross-references: GB:M86744; NID:g151550; PIDN:AAA55989.1; PID:g151552
 A:Experimental source: ATCC 19151
 C:Genetics:
 A:Gene: sdsA
 C:Keywords: sulfuric ester hydrolase

Query Match 49.4%; Score 39; DB 2; Length 528;
 Best Local Similarity 54.5%; Pred. No. 38;
 Matches 6; Conservative 4; Indels 1; Gaps 0;

OY 4 KUKMPYIAGY 14
 DB 257 KVKLPYIAGY 267

RESULT 13

T16609
 hypothetical protein K10B3.6 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
 C:Accession: T16609
 R:Gatung, S.
 submitted to the EMBL Data Library, February 1996
 A:Description: The sequence of C. elegans cosmid K10B3.
 A:Reference number: 218546
 A:Accession: T16609
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-757 <GAT>
 A:Cross-references: EMBL:U49941; NID:g1206038; PID:g1206045; PIDN:AAB53873.1; GSPDB:GN00
 A:Experimental source: strain Bristol N2; clone K10B3
 C:Genetics:
 A:Gene: CESP.K10B3.6
 A:Map position: X
 A:Intons: 89/2; 142/2; 184/2; 217/1; 269/3; 292/3; 330/3; 377/1; 425/2; 456/3; 497/3; 5

Query Match 49.4%; Score 39; DB 2; Length 757;
 Best Local Similarity 54.5%; Pred. No. 56;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 NUKKMPYI 11
 DB 558 INIKYPMY 568

RESULT 14

S48927
 hypothetical protein YHL041w - Yeast (*Saccharomyces cerevisiae*)
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002
 C:Accession: S48927
 R:Favella, T.
 submitted to the EMBL Data Library, June 1994
 A:Description: The sequence of S. cerevisiae cosmid 9196.
 A:Reference number: S46794
 A:Accession: S48927
 A:Molecule type: DNA
 A:Residues: 1-149 <FAV>
 A:Cross-references: EMBL:U11583; NID:g2289854; PID:g2289870; GSPDB:GN00008; MIPS:YHL041w
 C:Genetics:
 A:Gene: MIPS.YHL041w
 A:Cross-references: SGD:S0001033
 A:Map position: 8L
 C:Superfamily: *Saccharomyces* hypothetical protein YHL041w

Query Match 48.1%; Score 38; DB 2; Length 149;
 Best Local Similarity 54.5%; Pred. No. 15;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 KUKMPYIAGY 14

DB 55 KIKTPEIYIAGY 65

RESULT 15

A83651
 hypothetical protein BH0009 [imported] - *Bacillus halodurans* (strain C-125)
 C:Species: *Bacillus halodurans*
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: A83651
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: A83651
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-268 <STO>
 A:Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BA03728.1; GSPDB:GN
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH0009

Query Match 48.1%; Score 38; DB 2; Length 268;
 Best Local Similarity 53.8%; Pred. No. 28;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 NIKKMPYIAGY 14
 DB 162 SIQKPEAVVAGY 174

Search completed: April 20, 2003, 13:15:42
 Job time: 8.0785 secs

GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein & protein search, using sw model

Run on: April 20, 2003, 12:54:25; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524d-30

Sequence: 1 MNILKMPMTAGYK 15

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	374	1 SBP_CRYUA	P18632 cryptomeria
2	71	89.9	346	1 MPAL_CUPAR	Q9639 cupressus a
3	70	88.6	367	1 MPAL_JUNAS	P81234 juniperus a
4	59	74.7	375	1 MPAL_CHAOB	Q96385 chamaecypar
5	40	50.6	627	1 K2C1_MOUSE	P04104 mus musculu
6	39	49.4	347	1 RH31_YEAST	Q06624 saccharomyc
7	38	48.1	149	1 YH81_YEAST	P38730 caenorhabdi
8	38	48.1	614	1 YH81_YEAST	Q09320 caenorhabdi
9	38	48.1	614	1 YH81_YEAST	P28993 equine herp
10	38	48.1	166	1 K2C5_BOVIN	P04262 bos taurus
11	37	46.8	182	1 K2C4_BOVIN	P04260 bos taurus
12	37	46.8	182	1 K2C4_BOVIN	P04263 bos taurus
13	37	46.8	202	1 Y063_HUMAN	Q15040 homo sapien
14	37	46.8	370	1 K2C8_BOVIN	P05786 bos taurus
15	37	46.8	384	1 K2C8_HUMAN	P05787 homo sapien
16	37	46.8	482	1 K2C8_HUMAN	P05787 homo sapien
17	37	46.8	482	1 K2C8_HUMAN	P05787 homo sapien
18	37	46.8	493	1 VA18_VARV	Q10758 rattus norv
19	37	46.8	502	1 K2C8_XENLA	P33627 xenopus lae
20	37	46.8	511	1 CEST_HUMAN	Q08776 xenopus lae
21	37	46.8	511	1 NPE0_LOLPE	Q01440 loligo peal
22	37	46.8	512	1 K2C5_XENLA	P16878 xenopus lae
23	37	46.8	520	1 ION3_CARAU	P18520 carassius lae
24	37	46.8	524	1 K2C4_MOUSE	P07744 mus musculu
25	37	46.8	552	1 K2C6_MOUSE	P50446 mus musculu
26	37	46.8	563	1 K2C6_MOUSE	P02538 homo sapien
27	37	46.8	563	1 K2C6_MOUSE	P02538 homo sapien
28	37	46.8	563	1 K2C6_MOUSE	P02538 homo sapien
29	37	46.8	563	1 K2C6_MOUSE	P02538 homo sapien
30	37	46.8	563	1 K2C6_MOUSE	P02538 homo sapien
31	37	46.8	563	1 K2C6_MOUSE	P02538 homo sapien
32	37	46.8	621	1 VP40_HSVBC	Q01241 loligo peal
33	37	46.8	629	1 K2C3_HUMAN	P54817 bovine herp
					P12035 homo sapien

34	37	46.8	638	1 K220_HUMAN	Q01546 homo sapien
35	37	46.8	645	1 K22E_HUMAN	P35908 homo sapien
36	37	46.8	1203	1 Y741_CAEEL	Q11069 caenorhabdi
37	37	46.8	2214	1 POLG_CXAZ4	P36290 c genome po
38	36	45.6	188	1 EFP_FICPR	Q92677 rickettsia
39	36	45.6	273	1 YFCO_ECOLI	P76498 escherichia
40	36	45.6	358	1 Y993_METUA	Q58400 methanococc
41	36	45.6	389	1 TPEA_SULSH	Q05208 sulfolobus
42	36	45.6	422	1 K1CR_MOUSE	P05783 homo sapien
43	36	45.6	429	1 K1CR_MOUSE	P23233 xenopus lae
44	36	45.6	458	1 DESM_XENLA	P57002 neisseria m
45	36	45.6	460	1 PGMU_XENLA	

ALIGNMENTS

RESULT 1	ID	SBP_CRYUA	STANDARD	PRT	374 AA.
AC	P18632				
DI	01-NOV-1990 (Rel. 16, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	Sugi basic protein precursor (SBP) (Major allergen Cry j 1) (Cry j I).				
OS	Cryptomeria japonica (Japanese cedar).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.				
OX	NCBI_TaxID=3369;				
ON	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE=Pollen;				
RC	MEDLINE=94183234; PubMed=8135802;				
RA	Sone T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K.,				
RT	Kino K.,				
RT	"Cloning and sequencing of cDNA coding for Cry j I, a major allergen				
RT	of Japanese cedar pollen.";				
RL	Biochem. Biophys. Res. Commun. 199;619-625(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Pollen;				
RC	Namba M., Kurose M., Torigoe K., Fukuda S., Kurimoto M.;				
RL	Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.				
RP	SEQUENCE OF 22-41.				
RC	TISSUE=Pollen;				
RC	MEDLINE=89031257; PubMed=3181436;				
RA	Tanai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inouye S.,				
RT	Matunasi T.;				
RT	"N-terminal amino acid sequence of a major allergen of Japanese cedar				
RT	pollen (Cry j I).";				
RL	FEBS Lett. 239;329-332(1988).				
RP	[4]				
RC	CARBOHYDRATES.				
RC	TISSUE=Pollen;				
RC	MEDLINE=95003748; PubMed=7920021;				
RA	Hijikata A., Matsumoto I., Kojima K., Ogawa H.;				
RT	"Antigenicity of the oligosaccharide moiety of the Japanese cedar				
RT	(Cryptomeria japonica) pollen allergen, Cry j I.";				
RL	Int. Arch. Allergy Immunol. 105;198-202(1994).				
RN	[5]				
RP	STRUCTURE OF CARBOHYDRATES.				
RC	TISSUE=Pollen;				
RC	MEDLINE=95332249; PubMed=7608114;				
RA	Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Usui M.,				
RT	Fukuda S., Hanzawa H., Hanyama H., Kurimoto M.;				
RT	"Carbohydrate structures of the glycoprotein allergen Cry j I from				
RT	Japanese cedar (Cryptomeria japonica) pollen.";				
RL	J. Biochem. 117;289-295(1995).				
CC	-1- PTM: CONTAINS FUCCOSE/XULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.				
CC	-1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR				
CC	POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN.				
CC	-1- MISCELLANEOUS: THE SEQUENCE OF CRY J I FORM A IS SHOWN HERE. FORM				

B DIFFERS IN SIX POSITIONS.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A II/CRY J I SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D26544; BAA05542.1; -
CC EMBL; D26545; BAA05543.1; -
CC EMBL; D34639; BAA07020.1; -
CC PIR; A44773; A44773.
CC DR GlycoSiteDB; P18632; -
CC DR InterPro; IPR002022; Amb_allergen.
CC Pfam; PF00544; pec_lyase; 1.
CC PRINTS; PR00807; AMBALLERGEN.
CC Allergen; Glycoprotein; Multigene family; signal.
KW SIGNAL
FT 1 21
FT CHAIN 22 374
FT CARBOHYD 158 158
FT CARBOHYD 191 191
FT CARBOHYD 293 293
FT CARBOHYD 354 354
FT VARIANT 12 12
FT VARIANT 143 143
FT VARIANT 202 202
FT VARIANT 221 221
FT VARIANT 358 358
FT VARIANT 361 361
SQ SEQUENCE 374 AA; 40645 MW; 74AB25950248F56F CRC64;
Query Match 100.0%; Score 79; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. No. 2,1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNITKMPYIAGYK 15
DB 97 MNITKMPYIAGYK 111
RESULT 2
MEAL_CUPAR STANDARD; PRT; 346 AA.
ID MPAL_CUPAR
AC Q9SCG9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Cup a 1.
OS Cupressus arizonica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxId=99011;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=70571526; PubMed=11122214;
RA Acetuno E., Del Pozo V., Minguez A., Arieta I., Cortegano I.,
RA Cardaba B., Galarza S., Rojo M., Palomino P., Lahoz C.;
RT "Molecular cloning of major allergen from Cupressus arizonica pollen:
RT Cup a 1.";
RL Clin. Exp. Allergy 30:1750-1758(2000).
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY J I SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ243570; CAB62551.1; -
CC InterPro; IPR002022; Amb_allergen.
CC Pfam; PF00544; pec_lyase; 1.
CC PRINTS; PR00807; AMBALLERGEN.
KW Allergen; Glycoprotein.
FT CARBOHYD 127 127
FT CARBOHYD 157 157
FT CARBOHYD 272 272
SQ SEQUENCE 346 AA; 37589 MW; F1281DCDAD1D5DFD0 CRC64;
Query Match 89.9%; Score 71; DB 1; Length 346;
Best Local Similarity 80.0%; Pred. No. 6.3e-06;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNITKMPYIAGYK 15
DB 76 MNITKMPYIAGYK 90
RESULT 3
MEAL_JUNAS STANDARD; PRT; 367 AA.
ID MPAL_JUNAS
AC P81254; Q9ZNU7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Jun a 1 precursor.
OS Juniperus ashei (Ozark white cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Cupressaceae; Juniperus.
OX NCBI_TaxId=1101;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-38; 42-50; 58-80; 88-94;
RP 117-124; 134-140; 160-164; 256-263 AND 322-325.
RX TISSUE=Pollen;
RA MEDLINE=99414163; PubMed=10482836;
RA Midoro-Horiuchi T.M., Goldblum R.M., Kurosky A., Wood T.G.,
RA Brooks E.G.;
RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
RT allergen, Jun a 1.";
RL J. Allergy Clin. Immunol. 104:613-617(1999).
CC -1- DISSEAS: THIS PROTEIN CAUSES SEVERAL SEASONAL ALLERGIC RHINITIS
CC IN NORTH AMERICA.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY J I SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF106663; AAD03609.1; -
CC EMBL; AF106662; AAD03608.1; -
CC InterPro; IPR002022; Amb_allergen.
CC Pfam; PF00544; pec_lyase; 1.
CC PRINTS; PR00807; AMBALLERGEN.
KW Allergen; Glycoprotein; signal.

```

FT SIGNAL 1 21
FT CHAIN 22 367 MAJOR POLLEN ALLERGEN JUN A 1.
FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 367 AA; FC9B81E675662E49 CRC64;

Query Match
Best Local Similarity 88.6%; Score 70; DB 1; Length 367;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WNIXKMPYIAGYK 15
Db 97 WNIXKMPYIAGYK 111

RESULT 4
MPAL CHAOB STANDARD; PRT; 375 AA.
AC 096385;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Major pollen allergen Cha o 1 precursor.
OS Chamaecyparis obtusa (Japanese cypress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
OC Chamaecyparids.
OC NCBI_TaxID=13415;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=96265194; Pubmed=8676896;
RA Suzuki M., Komiyama N., Itoh H., Sone T., Kuno K., Takagi I.,
RT "Purification, characterization and molecular cloning of Cha o 1, a
RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen.";
RL Mol. Immunol. 33:451-460(1996).
CC -1- SIMILARITY: BELONGS TO THE POLYACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY J I SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Da5404; BAA08246.1; -
CC InterPro: IPR002022; Amb allergen.
CC DR Pfam: PF00544; pec lyase; 1.
CC DR PRINTS: PRO0807; AMBALLERGEN.
CC KW Allergen; Glycoprotein; Signal.
CC FT SIGNAL 1 21
CC FT CHAIN 22 375 MAJOR POLLEN ALLERGEN CHA O 1.
CC FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF706DBBF CRC64;

Query Match
Best Local Similarity 74.7%; Score 59; DB 1; Length 375;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WNIXKMPYIAGYK 15
Db 97 WNIXKMPYIAGYK 111

RESULT 5
K2C1_MOUSE
ID K2C1_MOUSE STANDARD; PRT; 627 AA.
AC P04104;
DT 01-NOV-1986 (Rel. 03, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal I (CytoKeratin I) (67 kDa cytoKeratin).
GN KRT1 OR KRT2-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 83-628 FROM N.A.
RX MEDLINE=85207740; Pubmed=2581964;
RA Steiner P.M., Parry D.A.D., Idler W.W., Johnson L.D., Steven A.C.,
RA Roop D.R.;
RT "Amino acid sequences of mouse and human epidermal type II keratins
RT of Mr 67,000 provide a systematic basis for the structural and
RT functional diversity of the end domains of Keratin Intermediate
RT filament subunits.";
RL J. Biol. Chem. 260:7142-7149(1985).
RN [2]
RP REVISIONS, AND SEQUENCE FROM N.A.
RA Roop D.R.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: HETEROHEPTAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC MICROFILAMENTAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M10937; AAD05191.1; -
CC PIR: A02951; KMS2.
CC DR SWISS-2DPAGE; P04104; MOUSE.
CC DR MGI: MGI:96698; Krt2-1.
CC DR InterPro: IPR001664; IF.
CC DR Pfam: PF00038; filament; 1.
CC DR PROSITE: PS00226; IF; 1.
CC KW Intermediate filament; Coiled coil; Keratin.
CC FT INIT MET 0 0
CC FT DOMAIN 1 186 HEAD.
CC FT DOMAIN 187 496 ROD.
CC FT DOMAIN 497 627 TAIL.
CC FT DOMAIN 187 222 COIL 1A.
CC FT DOMAIN 223 242 LINKER 1.
CC FT DOMAIN 243 333 COIL 1B.
CC FT DOMAIN 334 357 LINKER 12.
CC FT DOMAIN 358 496 COIL 2.
CC FT SITE 451 451 STUTTER.
CC SQ SEQUENCE 627 AA; 65092 MW; E7E848654539578 CRC64;

Query Match
Best Local Similarity 50.6%; Score 40; DB 1; Length 627;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 WNIXKMPYIAGYK 15
Db 476 WNIXKMPYIAGYK 490

RESULT 6
RH31_YEAST
ID RH31_YEAST STANDARD; PRT; 347 AA.
AC Q06624;
DT 01-NOV-1997 (Rel. 35, Created)

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA damage tolerance protein RHC31 (RAD31 homolog)
 GN AOS1 OR RHC31 OR YPR180W OR P9705.5
 OS Saccharomyces cerevisiae (Baker's yeast)
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 NCBI_TaxID=4932;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RX MEDLINE=97333271; PubMed=9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansgore W.,
 RA Aasey R., Aparicio A., Barrett B.G., Badcock K., Barnes V.,
 RA Boeslein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Deltis H., DiPaolo J., Dubois E., Duetchsch A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
 RA Hunkle-Smith S., Hyman R., Johnston M., Kaiman S., Klein K.,
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Matthe R., Messerguy F., Mewes H.-W., Mitrapati S., Moestl D.,
 RA Mueller-Puer S., Namath A., Nentwich U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Punelle D., Schaefer M., Scharfe M.,
 RA Scherens B., Schramm S., Schroeder W., Sclou A.M., Tettelin H.,
 RA Urestrazaru L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
 RA Walsh S.W., Wambutt R., Wang Y., Wedler H., Wilmott E.,
 RA Zhong M.V., Zollner A., Vo D.H., Han J.
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI."
 RL Nature 387:103-105(1997).
 RN
 RP CHARACTERIZATION.
 RX MEDLINE=9724684; PubMed=9092625;
 RA Shavgedi M., Does C.L., Tavaresoli M., Watts F.Z.
 RT "Characterisation of Schizosaccharomyces pombe rad31, a UBA-related
 RT gene required for DNA damage tolerance."
 RL Nucleic Acids Res. 25:1162-1169(1997).
 CC -1- FUNCTION: COULD BE INVOLVED IN A UBIQUITIN-RELATED PROCESS
 CC -1- IMPORTANT: FOR DNA DAMAGE TOLERANCE.
 CC -1- SIMILARITY: TO THE N-TERMINAL OF UBIQUITIN-ACTIVATING ENZYME E1.
 CC -1- SIMILARITY: STRONG, TO S.POMBE RAD31.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL, U25842; AAB68113.1; -
 DR SGD; S0006384; AOS1.
 DR InterPro; IPR000594; Thif_domain.
 DR Pfam; PF00899; Thif; 1.
 KW DNA damage.
 SQ SEQUENCE 347 AA; 39273 MW; 1ADB7B817BA27F3E CRC64;
 QY 4 KLMMPYIAG 13
 DB 145 KLMPIYVAG 154
 ID YH01-1997
 AC YH01-1997
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)

DE Hypothetical 17.6 kDa protein in CBP2 5' region.
 GN YH041W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 NCBI_TaxID=4932;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Lacroille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
 RA Nham N., Rifkin L., Riley L., St Peter H., Trevisan E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterson R., Wilson R.,
 RA Vaadin M.
 RA "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VIII."
 RT Science 265:2077-2082(1994).
 RL
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL, U11583; AAB65053.1; -
 DR PIR; S48927; S48927.
 DR SGD; S0001033; YH041W.
 KW Hypothetical protein.
 SQ SEQUENCE 149 AA; 17554 MW; 5DE657D38A83CFDF CRC64;
 QY 4 KLMMPYIAG 14
 DB 55 KLMPIYVAG 65
 ID YH02-1997
 AC YH02-1997
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 69.0 kDa protein F40B5.2 in chromosome X.
 GN F40B5.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Geisel C.
 RA Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RL -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: TO C. ELEGANS T23G7.2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

```

DR EMBL; U23182; AAA64337.1; -
DR Wormpep; F40B5.2; CE01936.
DR InterPro; IPR004843; M-peptidase.
DR Pfam; PF00149; Metallophos. 1.
KW Hypothetical protein; transmembrane.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 129 149 POTENTIAL.
FT TRANSMEM 156 176 POTENTIAL.
FT TRANSMEM 205 225 POTENTIAL.
FT TRANSMEM 416 436 POTENTIAL.
SQ SEQUENCE 614 AA; 68990 MW; 20BBB8AB524EDDB CRC64;

Query Match
Best Local Similarity 48.1%; Score 38; DB 1; Length 614;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 MWIKLMPYIAGY 14
DB 15 MWIKLKIATIVIGF 28

RESULT 9
VP40_HSVEB STANDARD; PRT; 646 AA.
AC P28936; O69263;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Capsid protein VP40 (Virion structural gene 35 protein) [Contains:
DE Capsid protein VP24 (Assemblin) (Protease) (EC 3.4.21.97); Capsid
DE protein VP22a].
GN 35.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=31520;
RN NCB1
RP SEQUENCE FROM N.A.
RX MEDLINE=92395566; PubMed=318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1."
RT Virology 189:304-316(1992).
CC -1- FUNCTION: VP22A IS A COMPONENT OF THE CAPSID CORE INVOLVED IN
CC PROCESSING AND PACKAGING OF PROGENY DNA. VP24 IS A PROTEASE
CC WHICH CAN PROTEOLYTICALLY CLEAVE ITSELF AND VP22A AT THE C-
CC TERMINUS.
CC -1- CATALYTIC ACTIVITY: Cleaves -Ala-Ser- and -Ala-Ala-bonds in
CC the scaffold protein.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S21.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP40.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M86664; AAB02470.1; -
DR EMBL; M86664; AAB02471.1; -
DR PIR; I36798; WZBEC8.
DR HSSP; P16753; WZBEC8.
DR MEROPS; S21.001; -
DR InterPro; IPR001847; Assemblin.
DR Pfam; PF00716; Peptidase S21, 1.
DR PRINTS; PR000216; HSVCAPSIDP40.
KW Coat protein; Hydrolase; Serine protease.
FT CHAIN 1 GENE 35 PROTEIN.
FT CHAIN 1 646 GENE 35.5 PROTEIN.
FT CHAIN 1 242 COAT PROTEIN VP24 (PROTEASE).
FT CHAIN 243 7622 COAT PROTEIN VP22A.

```

```

FT PROPEP 7623 646 C-TERMINAL PEPTIDE.
FT SITE 242 243 CLEAVAGE (BY THE PROTEASE) (PROBABLE).
FT SITE 622 623 CLEAVAGE (BY THE PROTEASE) (PROBABLE).
FT ACT_SITE 55 55 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 142 142 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 646 AA; 68579 MW; 1C1C28C848116276 CRC64;

Query Match
Best Local Similarity 48.1%; Score 38; DB 1; Length 646;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LKMPYIAGY 14
DB 11 VSLPIYAGY 20

RESULT 10
K2C5_BOVIN STANDARD; PRT; 166 AA.
ID K2C5_BOVIN
AC P04262;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal 68 kDa, component IB (fragment).
OS Bos taurus (Bovine).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Bovidae; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OX NCBI_TaxID=9913;
RN NCB1
RP SEQUENCE FROM N.A.
RX MEDLINE=85128114; PubMed=6084625;
RA Jorcano J.L., Franz J.K., Franke W.W.;
RT "Amino acid sequence diversity between bovine epidermal cyokeratin
RT polypeptides of the basic (type II) subfamily as determined from cDNA
RT clones."
RL Differential 28:155-163(1984).
CC -1- SUBUNIT: HETEROTRIMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC MICROFIBRILLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; K03534; AAA30601.1; -
DR PIR; A02948; KRB02B.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; Filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Keratin.
FT NON_TER 1
FT DOMAIN 1 ROD.
FT DOMAIN 42 166 TAIL.
FT DOMAIN 41 COIL 2B.
SQ SEQUENCE 166 AA; 15505 MW; 4BPB495A7C6B4B0 CRC64;

Query Match
Best Local Similarity 46.8%; Score 37; DB 1; Length 166;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MWIKLMPYIAGYK 15
DB 17 MWIKLADIVETATYR 31

```



```

RESULT 11
K2C4 BOVIN STANDARD; PRT; 182 AA.
AC P04260;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal 55 kDa, component IV (Keratin K6 gamma)
DE (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85128114; PubMed=6084625;
RA Jorcano J.L., Franz J.K., Franke W.W.;
RT "Amino acid sequence diversity between bovine epidermal cytoke-
RT polypeptides of the basic (type II) subfamily as determined from cDNA
RT clones."
RL Differentiation 28:155-163(1984).
RN [2]
RP SEQUENCE OF 99-182 FROM N.A.
RA Navarro M., Segreles C., Jorcano J.L.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: HETEROPTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC MICROFILAMENTAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03536; AAA30603.1; -
DR EMBL; Z48499; CA88399.1; -
DR PIR; A02946; A02946.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Keratin.
FT NON TER 1 1
FT DOMAIN 1 83 ROD.
FT DOMAIN 84 182 TAIL.
FT DOMAIN <1 83 COIL 2.
FT SITE 27 27 STUTTER.
SQ SEQUENCE 182 AA; 18670 MW; 740AF9906AE3B66C CRC64;

Query Match 46.8%; Score 37; DB 1; Length 182;
Best Local Similarity 46.7%; Pred. No. 9.1;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 MNKIKMPYIAGYK 15
DB 63 MNVKLADVEIATYR 77

```

```

OY 1 MNKIKMPYIAGYK 15
DB 42 MNVKLADVEIATYR 56

RESULT 13
Y063 HUMAN STANDARD; PRT; 202 AA.
AC Q15040;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA0063 (HA1234).
GN KIAA0063.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=bone marrow; PubMed=7584044;
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kawarabayashi Y., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1."
RL DNA Res. 1:223-229(1994).

```


[2]
 SEQUENCE FROM N.A.
 MEDLINE=20057165; PubMed=10591208;
 Dunham I., Hunt A.R., Collins J.E.,
 Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 Bagunley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 Evans K.L., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
 Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 Laird G.K., Langford C.F., Leverisha M.A., Lloyd A.M., Lloyd D.M.,
 Matyln I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
 Mcclay J., McLaren S., Mcmurray A.A., Milne S.L., Mott B.J.C.T.,
 Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
 Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 Soederlund C., Spragon L., Steward C.A., Suleston J.E., Swann R.M.,
 Vandin M., Wallis J.M., Whiteley M.N., Wilmer T.E., Wilting L.,
 Williams L., Williams S.A., Williamson H., Wiltner D.L.,
 Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,
 Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 Dorman A., Fang F., Fu Y., Hu A., Kenyon S., Lai H., Lao H.I.,
 Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 Phan S., Qi S., Qian Y., Ray L., Ren Q., Shauli S., Sloan D., Song L.,
 Wang Q., Wang Y., Wang Z., White J., Williams D., Wu H., Yao Z.,
 Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
 Fulton R., Johnson D., Bems G., Bentley D., Bradshaw H., Bourne S.,
 Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 Hands K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
 Schuet P., Walker C., Wamsley J., Wamsley P., Bepin K., Nelson J.,
 Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
 Emanuel B.S., Shaich T., Kuzanishi H., Salta S., Budarf M.L.,
 Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
 Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyraud M., Kedra D.,
 Seroussi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P.,
 Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
 Tiliun Y., Wright H.,
 "The DNA sequence of human chromosome 22."
 Nature 402:489-495(1999).
 [3]
 SEQUENCE FROM N.A.
 TISSUE=Skin;
 Strausberg R.;
 Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

Query Match 46.8%; Score 37; DB 1; Length 202;
 Best Local Similarity 53.8%; Pred. No. 10;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNKIKMPYIAGX 13
 DB 153 LPSKCKMPWIGG 165

RESULT 14

K2C8 BOVIN

ID K2C8_BOVIN STANDARD; PRT; 370 AA.

AC P05786;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Keratin, type II cytoskeletal 8 (Cytokeratin 8) (Cytokeratin A)

DE (Fragment)

OS Bos taurus (bovine)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

[1]

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

Query Match 46.8%; Score 37; DB 1; Length 370;
 Best Local Similarity 46.7%; Pred. No. 19;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

QY 1 MNKIKMPYIAGX 15
 DB 277 MNVKAIDVETIATYR 291

DT 01-FEB-1996 (Rel. 33, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Keratin, type II cytoskeletal 6D (Cytokeratin 6D) (CK 6D) (K6D
 DE Keratin) (Fragment).
 GN KRT6D.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=95355491; PubMed=7543104;
 RA Takahashi K., Paladini R.D., Coulombe P.A.;
 RT "Cloning and characterization of multiple human genes and cDNAs
 encoding highly related type II keratin 6 isoforms."
 RL J. Biol. Chem. 270:18581-18592(1995).
 CC -1- FUNCTION: THERE ARE TWO TYPES OF CYTOSKELETAL AND MICROFILLLAR
 CC KERATIN: I (ACIDIC; 40-55 kDa) [K9 TO K20] AND II (NEUTRAL TO
 CC BASIC; 56-70 kDa) [K1 TO K8]. BOTH A BASIC AND AN ACIDIC KERATIN
 CC ARE REQUIRED FOR FILAMENT ASSEMBLY.
 CC -1- SUBUNIT: HETERODIMER OF A TYPE I AND A TYPE II KERATIN. KRT6
 CC ISOMERS ASSOCIATE WITH KRT16 AND/OR KRT17.
 CC -1- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN DISTINCT TYPES
 CC OF EPITHELIA SUCH AS THOSE IN ORAL MUCOSA, ESOPHAGUS, PAPILLAE
 CC OF TONGUE AND HAIR FOLLICLE OUTER ROOT SHEATH.
 CC -1- MISCELLANEOUS: THERE ARE AT LEAST SIX ISOFORMS OF HUMAN TYPE II
 CC KERATIN 6 (K6), K6A BEING THE MOST ABUNDANT REPRESENTING ABOUT 77%
 CC OF ALL FORMS FOUND IN EPITHELIA.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; L42610; AAB60696.1; -
 DR EMBL; L42603; AAB60696.1; JOINED.
 DR EMBL; L42604; AAB60696.1; JOINED.
 DR EMBL; L42605; AAB60696.1; JOINED.
 DR EMBL; L42606; AAB60696.1; JOINED.
 DR EMBL; L42607; AAB60696.1; JOINED.
 DR EMBL; L42608; AAB60696.1; JOINED.
 DR EMBL; L42609; AAB60696.1; JOINED.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Keratin; Multigene family.
 FT NON_TER 1
 FT DOMAIN <1 292 ROD.
 FT DOMAIN 293 384 TAIL.
 FT DOMAIN <1 18 COIL_1A.
 FT DOMAIN 19 37 LINKER_1.
 FT DOMAIN 38 129 COIL_1B.
 FT DOMAIN 130 153 LINKER_12.
 FT DOMAIN 154 292 COIL_2.
 FT SITE 234 234 STUTTER.
 SQ SEQUENCE 384 AA; 42468 MW; D69DCFD808C94E91 CRC64;
 OY 1 MNILKMPYIAGYK 15
 DB 272 MNTVLADVEIATYR 286
 Query Match 46.8%; Score 37; DB 1; Length 384;
 Best Local Similarity 46.7%; Pred. No. 20;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 12.6711 Seconds

(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524d-30

Sequence: 1 MNIKLKMPTIAGYK 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	374	10	Q8RUR1
2	71	89.9	367	10	Q93XL6
3	70	88.6	367	10	Q93XL6
4	70	88.6	367	10	Q91L12
5	70	88.6	367	10	Q91L12
6	70	88.6	367	10	Q9M4S6
7	70	88.6	367	10	Q9M4S6
8	70	88.6	367	10	Q9M4S6
9	70	88.6	367	10	Q9M4S6
10	70	88.6	367	10	Q9M4S6
11	70	88.6	367	10	Q9M4S6
12	70	88.6	367	10	Q9M4S6
13	70	88.6	367	10	Q9M4S6
14	70	88.6	367	10	Q9M4S6
15	70	88.6	367	10	Q9M4S6
16	70	88.6	367	10	Q9M4S6

17	41	51.9	1017	5	Q9VM32	Q9VM32 drosophila
18	41	51.9	1030	5	Q9GZG7	Q9GZG7 drosophila
19	41	51.9	1086	5	Q9VM31	Q9VM31 drosophila
20	40	50.6	184	17	Q977C2	Q977C2 sulfolobus
21	40	50.6	330	17	Q8TL23	Q8TL23 methanococcus
22	40	50.6	357	5	Q62005	Q62005 branchiostoma
23	40	50.6	414	16	Q8XM91	Q8XM91 branchiostoma
24	40	50.6	522	10	Q9LFX0	Q9LFX0 arabidopsis
25	40	50.6	627	11	Q9D2K8	Q9D2K8 mus musculus
26	40	50.6	650	2	Q9S6S6	Q9S6S6 lactococcus
27	39	49.4	162	16	Q9JW3	Q9JW3 neisseria m
28	39	49.4	246	2	Q9XIA1	Q9XIA1 bacteroides
29	39	49.4	294	8	Q02676	Q02676 podospira a
30	39	49.4	350	5	Q62001	Q62001 branchiostoma
31	39	49.4	395	16	Q8ZAZ8	Q8ZAZ8 yeastina pe
32	39	49.4	528	2	Q52556	Q52556 pseudomonas
33	39	49.4	757	5	Q21407	Q21407 caenorhabdi
34	38	48.1	227	8	Q8WD75	Q8WD75 lepidosiren
35	38	48.1	268	16	Q9RC96	Q9RC96 bacillus ha
36	38	48.1	320	16	Q9CEA7	Q9CEA7 lactococcus
37	38	48.1	335	5	Q62000	Q62000 branchiostoma
38	38	48.1	365	5	Q61999	Q61999 branchiostoma
39	38	48.1	378	13	Q90WPO	Q90WPO trachemys s
40	38	48.1	444	5	Q9V639	Q9V639 drosophila
41	38	48.1	455	5	Q8ST16	Q8ST16 caenorhabdi
42	38	48.1	482	5	Q8ST17	Q8ST17 caenorhabdi
43	38	48.1	541	16	Q9KUB4	Q9KUB4 vibrio chol
44	38	48.1	551	4	Q9S678	Q9S678 homo sapien
45	38	48.1	551	4	Q9NSA9	Q9NSA9 homo sapien

ALIGNMENTS

Q8RUR1	1	PRELIMINARY;	PRT;	374 AA.
AC	Q8RUR1			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Cry j 1 precursor.			
GN	Cry j 1.1 OR Cry j 1.2.			
OS	Cryptomeria japonica (Japanese cedar).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.			
OX	NCBI_TaxID=3369;			
[1]				
RP	SEQUENCE FROM N.A.			
RC	TISSUE=POLLEN;			
RA	Futamura N., Shinohara K.,			
RT	"Isolation and characterization of cDNAs encoding major allergen Cry j			
RT	1 from Cryptomeria japonica pollen."			
RU	Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AB081309; BAB86286.1; -			
DR	EMBL; AB081310; BAB86287.1; -			
KW	Signal.			
FT	SIGNAL.			
FT	CHAIN			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT	CRY j 1			
FT	POTENTIAL.			
FT	CRY j 1			
FT	SEQUENCE			
FT	374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;			
FT</				

```

AC 093XL6;* (TREMBlrel. 19, Created).
DT 01-DEC-2001 (TREMBlrel. 19, Created).
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative allergen Cup a 1 precursor.
GN CUP A 1.
OS Cupressus arizonica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=49011;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=POLLEN;
RA Butcheri C., Di Felice G., Pini C.;
RT "Cloning of Cupressus arizonica major allergen.";
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL, AJ278498; CAC37790.2; -.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 367 AA; 39809 MW; AAF97260423A9F28 CRC64;

Query Match 89.9%; Score 71; DB 10; Length 367;
Best Local Similarity 80.0%; Pred. No. 5.6e-05;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNIKLMPYIAGYK 15
DB 97 MNIKLMPYVAGHK 111

RESULT 3
Q9LIT2 PRELIMINARY; PRT; 367 AA.
ID Q9LIT2;
AC Q9LIT2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-2.
OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX NCBI_TaxID=39584;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=2135424; PubMed=11422137;
RA Midoro-Horitsu T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778 (2001).
DR EMBL, AF151427; AAF80164.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF89586 CRC64;

Query Match 88.6%; Score 70; DB 10; Length 367;
Best Local Similarity 80.0%; Pred. No. 8.6e-05;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNIKLMPYIAGYK 15
DB 97 MNIKLMPYVAGHK 111

RESULT 4
Q9LIT1 PRELIMINARY; PRT; 367 AA.
ID Q9LIT1;
AC Q9LIT1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-1.

```

```

OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX NCBI_TaxID=39584;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=2135424; PubMed=11422137;
RA Midoro-Horitsu T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778 (2001).
DR EMBL, AF151429; AAF80166.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74F4711 CRC64;

Query Match 88.6%; Score 70; DB 10; Length 367;
Best Local Similarity 80.0%; Pred. No. 8.6e-05;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNIKLMPYIAGYK 15
DB 97 MNIKLMPYVAGHK 111

RESULT 5
Q9M4S6 PRELIMINARY; PRT; 367 AA.
ID Q9M4S6;
AC Q9M4S6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP s1.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RC Monnaie R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257491; AAF72625.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39835 MW; B4E9C60108C2C5A3 CRC64;

Query Match 88.6%; Score 70; DB 10; Length 367;
Best Local Similarity 80.0%; Pred. No. 8.6e-05;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNIKLMPYIAGYK 15
DB 97 MNIKLMPYVAGHK 111

RESULT 6
Q9M4S4 PRELIMINARY; PRT; 367 AA.
ID Q9M4S4;
AC Q9M4S4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP s1.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;

```

RN [1]
RP SEQUENCE FROM N.A.
RA Monesalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
of Cupressus sempervirens."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257493; AAF72627.1; -
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 88.6%; Score 70; DB 10; Length 367;
Best Local Similarity 80.0%; Pred. No. 8.6e-05;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNIKLKMPLYAGYK 15
DB 97 MNIKLKMPLYAGHK 111

RESULT 7

ID Q9M4S3 PRELIMINARY; PRT; 367 AA.
AC Q9M4S3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUPSL.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxId=31469;

RA [1]
RP SEQUENCE FROM N.A.
RA Monesalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
of Cupressus sempervirens."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257494; AAF72628.1; -
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39832 MW; BSDPBF5A61C07A53 CRC64;

Query Match 88.6%; Score 70; DB 10; Length 367;
Best Local Similarity 80.0%; Pred. No. 8.6e-05;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNIKLKMPLYAGYK 15
DB 97 MNIKLKMPLYAGHK 111

RESULT 8

ID Q9M4S2 PRELIMINARY; PRT; 367 AA.
AC Q9M4S2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUPSL.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxId=31469;

RN [1]
RP SEQUENCE FROM N.A.
RA Monesalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
of Cupressus sempervirens."

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257495; AAF72629.1; -
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39819 MW; AF7E055A61C07A53 CRC64;

Query Match 88.6%; Score 70; DB 10; Length 367;
Best Local Similarity 80.0%; Pred. No. 8.6e-05;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNIKLKMPLYAGYK 15
DB 97 MNIKLKMPLYAGHK 111

RESULT 9

ID Q93X51 PRELIMINARY; PRT; 367 AA.
AC Q93X51;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative allergen jun o 1.
GN JUN O 1.
OS Juniperus oxycedrus (Prickly juniper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX NCBI_TaxId=69008;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=POLLEN;
RA Iacovacci P., Di Felice G., Pini C.;
RT "Cloning of Juniperus oxycedrus major allergen."
RL Submitted (AVG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ933767; CAC48400.1; -
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
SQ SEQUENCE 367 AA; 39808 MW; SD28204DBFD1B9D7 CRC64;

Query Match 88.6%; Score 70; DB 10; Length 367;
Best Local Similarity 80.0%; Pred. No. 8.6e-05;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNIKLKMPLYAGYK 15
DB 97 MNIKLKMPLYAGHK 111

RESULT 10

ID Q9M4S5 PRELIMINARY; PRT; 367 AA.
AC Q9M4S5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUPSL.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxId=31469;

RN [1]
RP SEQUENCE FROM N.A.
RA Monesalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
of Cupressus sempervirens."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257492; AAF72626.1; -
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0E3263B741 CRC64;

Query Match 83.5%; Score 66; DB 10; Length 367;
 Best Local Similarity 73.3%; Pred. No. 0.00048;
 Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNKMKPMYIAGYK 15
 |||||:|:|:|:
 Db 97 MNKMKPMYIAGYK 111

RESULT 11

ID 021303 PRELIMINARY; PRT; 1202 AA.

AC 021303;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE K07G5.3 protein.
 GN K07G5.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 CX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurtry A.A.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RT [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99069613; PubMed=9851916;
 RX none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: 271264; CAA95830.1; -
 DR InterPro: IPR000008; C2.
 DR SMART; SM00239; C2; 1.
 DR SEQUENCE 1202 AA; 138469 MW; 6531124389BDD063 CRC64;

Query Match 55.7%; Score 44; DB 5; Length 1202;
 Best Local Similarity 53.8%; Pred. No. 20;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 IKKMPMYIAGYK 15
 |:|:|:|:|:
 Db 811 IRLNIPLYISNYK 823

RESULT 12

ID 023625 PRELIMINARY; PRT; 345 AA.

AC 023625;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE ZK829.8 protein.
 GN ZK829.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 CX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harris B.R.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RT [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99069613; PubMed=9851916;
 RX none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: 273899; CAA98078.1; -

DR InterPro: IPR003002; 7TM_Chemol.
 DR InterPro: IPR000168; 7TM_nematode.
 DR Pfam: PF01461; 7tm_4; 1.
 DR SEQUENCE 345 AA; 39398 MW; 025C9DBF14FD5299 CRC64;

Query Match 54.4%; Score 43; DB 5; Length 345;
 Best Local Similarity 53.8%; Pred. No. 8.9;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 IKKMPMYIAGYK 15
 :|:|:|:|:
 Db 30 IKKSPRYIGYR 42

RESULT 13

ID 0941V8 PRELIMINARY; PRT; 645 AA.

AC 0941V8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Hypothetical 72.5 KDa protein.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartioideae; Oryzaceae; Oryza.
 CX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CV. NIPONBARE;
 RC Buehl C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
 RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsilira T.,
 RA Riggs F., Hsiao J., Ziemann V., Blunt S., Pai G., Vanaken S.E.,
 RA Uterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
 RA White O., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJUBa0082M15 genomic sequence."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC020666; AAK43495.1; -
 DR InterPro: IPR003656; BBD_finger.
 DR InterPro: IPR002048; EF-hand.
 DR PROSITE; PS00019; EF_HAND; UNKNOWN_1.
 DR KW Hypothetical protein.
 DR SEQUENCE 645 AA; 72497 MW; B95683FB4F753C51 CRC64;

Query Match 54.4%; Score 43; DB 10; Length 645;
 Best Local Similarity 63.6%; Pred. No. 17;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 KKKMPMYIAGY 14
 |||||:|:|:|:
 Db 437 KKKPLMLAGY 447

RESULT 14

ID 0852E8 PRELIMINARY; PRT; 965 AA.

AC 0852E8;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
 DE P0022F10.3 protein.
 GN P0022F10.3.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartioideae; Oryzaceae; Oryza.
 CX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CV. NIPONBARE;
 RC Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa niponbare (GA3) genomic DNA, chromosome 1, PAC
 clone: P0022F10.";

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003229; BAB89488.1; -
 SQ SEQUENCE 965 AA; 109160 MW; F80F90584944694E CRC64;

Query Match 54.4%; Score 43; DB 10; Length 965;
 Best Local Similarity 63.6%; Pred. No. 25;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 KLMKMPYIAGY 14
 ||| ||| |||
 Db 696 KLTPLPLAGY 706

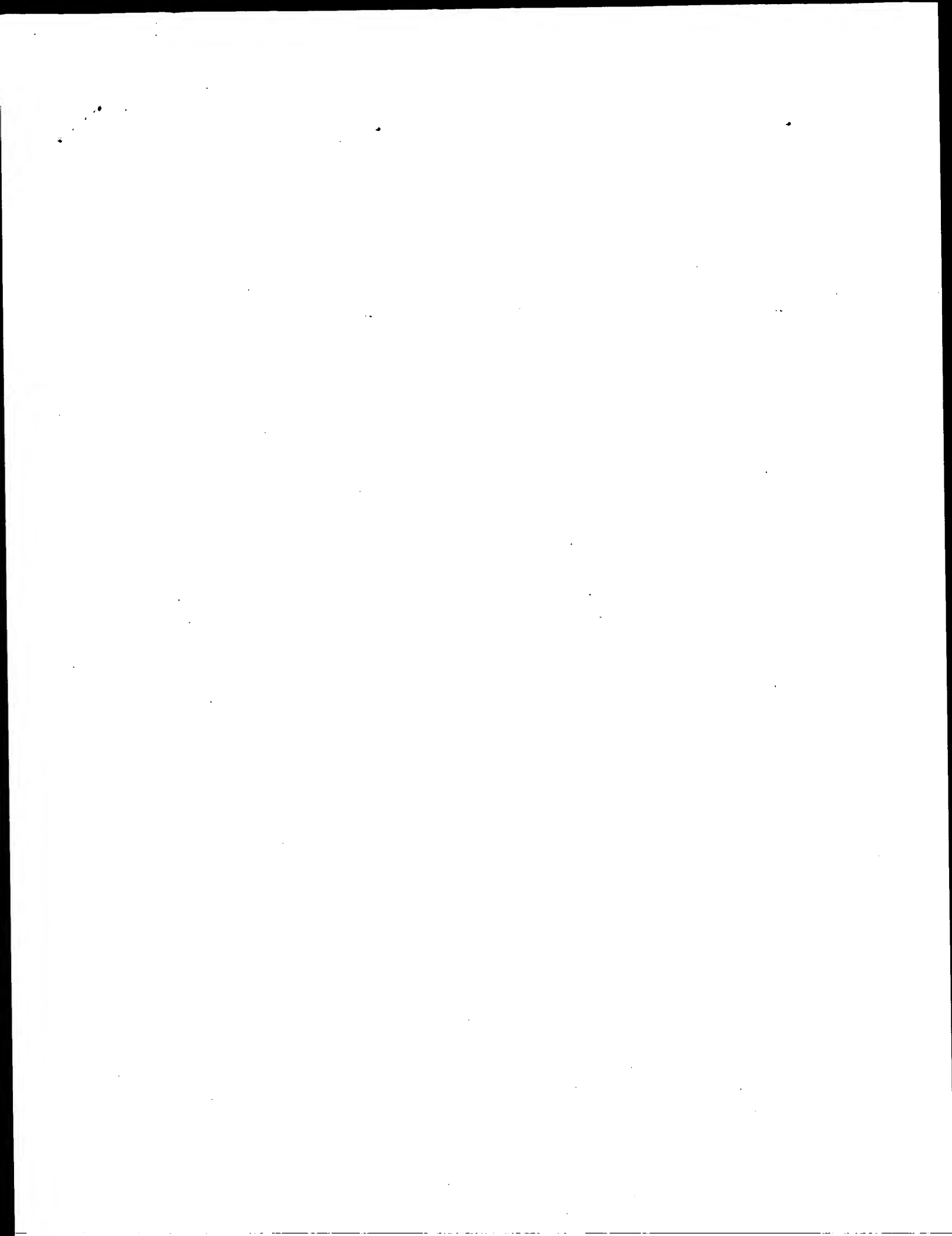
RESULT 15

062002 PRELIMINARY; PRT; 839 AA.
 ID 062002
 AC 062002;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Intermediate filament protein C1 (Fragment).
 OS Branchiostoma Filicidae (Florida lancelet) (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 OX NCBI_TaxID=7739;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98267218; PubMed=9602172;
 RA Riemer D., Karabinos A., Weber K.;
 RT "Analysis of eight cDNAs and six genes for intermediate filament (IF)
 RT proteins in the cephalochordate Branchiostoma reveals differences in
 RT the IF multigene families of lower chordates and the vertebrates."
 RL Gene 211361-373 (1998).
 CC -I- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 DR EMBL; AJ223577; CAAL1444.1; -
 DR InterPro; IPR01664; IF.
 DR Pfam; PF00038; filament; 6.
 DR PROSITE; PS00226; IF; 1.
 KM Coiled coil; Intermediate filament.
 FT NON TER 1
 SQ SEQUENCE 839 AA; 96156 MW; 902F6DB805CAFP4 CRC64;

Query Match 53.2%; Score 42; DB 5; Length 839;
 Best Local Similarity 53.3%; Pred. No. 33;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 MNILKMPYIAGYK 15
 ||| ||| |||
 Db 319 MNILKLPTEIAAYR 333

Search completed: April 20, 2003, 13:13:03
 Job time : 14.6711 secs



GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds
(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524d-31

Perfect score: 83

Sequence: 1 KMPMYIAGYKTFDGR 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : *

1: PIR.73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	100.0	374	2 JC2124	major allergen Cry j I precursor (clone PCCI-15) - Japanese cedar
2	83	100.0	374	2 JC2123	major allergen Cry j I precursor (clone PCCI-15) - Japanese cedar
3	48	57.8	157	2 H87601	transcription elon
4	46	55.4	477	2 S76496	hypothetical prote
5	44	53.0	357	2 E90116	mRNA capping enzym
6	44	53.0	379	2 A70384	glutamate N-acetyl
7	43	51.8	419	2 S76571	hypothetical prote
8	42	50.6	337	2 S26863	glyceroldehyde-3-P
9	42	50.6	337	2 S40610	glyceroldehyde-3-P
10	42	50.6	337	2 T30889	valine-tRNA ligase
11	41.5	50.0	1218	2 T44562	catalase (EC 1.11.1.1)
12	41	49.4	321	2 S49369	mobilitization prote
13	41	49.4	542	2 T06728	pectate lyase (EC 3.1.1.1)
14	41	49.4	952	2 S38653	transposase - Kleb
15	41	49.4	1071	2 T30881	glyceroldehyde-3-P
16	40	48.2	337	1 DEJ0GC	dynamin heavy chain
17	40	48.2	368	2 G86427	glyceroldehyde-3-P
18	40	48.2	974	2 A90140	conserved pectate 1
19	39	47.0	162	2 E91113	hydrogenase-2 oper
20	39	47.0	162	2 E85958	member of hyb oper
21	39	47.0	162	2 F65085	hydrogenase-2 oper
22	39	47.0	162	2 T39115	formamide-2 oper
23	39	47.0	162	2 B36936	nitrogenase (EC 1.1.1.1)
24	38	45.8	187	2 G89922	conserved hypothe
25	38	45.8	298	2 C41047	exonzyme S synthet
26	38	45.8	316	2 G72077	CT007 hypothetical
27	38	45.8	316	2 G86545	alcohol dehydrogen
28	38	45.8	335	2 AC3189	pectate lyase-like
29	38	45.8	341	2 T47653	pectate lyase-like

30	38	45.8	404	2 S12209	pectate lyase (EC 3.1.1.1)
31	38	45.8	404	2 A69467	hypothetical prote
32	38	45.8	418	2 T39058	probable isocitrat
33	38	45.8	529	2 A83517	probable aldehyde
34	38	45.8	568	2 T03950	probable glucose-6
35	38	45.8	605	2 T43191	probable pyruvate
36	38	45.8	688	2 H96881	protein P1E22.10 (
37	37.5	44.2	531	1 SYNCYT	tyrosine-tRNA liga
38	37	44.6	125	2 E59106	hypothetical prote
39	37	44.6	158	2 E97214	uncharacterized me
40	37	44.6	171	2 C64247	conserved hypothe
41	37	44.6	176	2 C70599	hypothetical prote
42	37	44.6	226	2 H70400	rare lipoprotein A
43	37	44.6	265	2 D82871	conserved hypothe
44	37	44.6	267	2 A75217	hypothetical prote
45	37	44.6	311	2 JC5943	catechol 1,2-dioxy

ALIGNMENTS

RESULT 1

major allergen Cry j I precursor (clone PCCI-15) - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C/Accession: JC2124

R/Source, T.: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese

A/Reference number: JC2123; MUID:94183234; PMID:8155802

A/Accession: JC2124

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-References: GB:026545; NID:9493633; PIDN:BAA05543.1; PID:9493634

A/Experimental source: pollen

A/Note: The authors described carbohydrate binding site for residue 279

C/Superfamily: pectate lyase LRT59

C/Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-374/Product: major allergen Cry j I (clone PCCI-15) #status predicted <MAT>

F:158-191,293,354/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 100.0%; Score 83; DB 2; Length 374;

Best Local Similarity 100.0%; Pred. No. 6.6e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMPMYIAGYKTFDGR 15

DB 102 KMPMYIAGYKTFDGR 116

RESULT 2

JC2123

major allergen Cry j I precursor (clone PCCI-2-2) - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C/Accession: JC2123; PC2065

R/Source, T.: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese

A/Reference number: JC2123; MUID:94183234; PMID:8155802

A/Accession: JC2123

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-References: GB:026544; NID:9493633; PIDN:BAA05542.1; PID:9493632

A/Experimental source: pollen

A/Note: The authors described carbohydrate binding site for residue 279

C/Superfamily: pectate lyase LRT59

C/Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-374/Product: major allergen C1y 1 (clone pCCT-2-2) #status predicted <MAT>
F:158,191,293,354/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match 100.0%; Score 83; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 6,6e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTFDGR 15
|||
Db 102 KMPMYIAGYKTFDGR 116

RESULT 3

Transcription elongation factor GreA CC2848 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C/Accession: H87601

R:Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Lamb, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kohn
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: H87601

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-157 <STO>

A/Cross-references: GB:AB005673; NID:g13424460; PIDN:AAK24812.1; GSPDB:GN00148

C/Genetics:
C/Superfamily: transcription elongation factor greB

C/Keywords: transcription factor

Query Match 57.8%; Score 48; DB 2; Length 157;
Best Local Similarity 61.5%; Pred. No. 0.46;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTFD 13
|||
Db 3 KMPMYIAGYQTLD 15

RESULT 4
S76496
Hypothetical protein - Synechocystis sp. (strain PCC 6803)

C/Species: Synechocystis sp.

A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C/Accession: S76496
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
sp.

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S76496

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-477 <KAN>

A/Cross-references: EMBL:U00915; GB:AB001339; NID:g1653604; PIDN:BA18625.1; PID:g165371

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Superfamily: starch synthase

Query Match 55.4%; Score 46; DB 2; Length 477;
Best Local Similarity 46.7%; Pred. No. 3.3;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTFDGR 15
|||
Db 79 KIPPLYFGHPAFDR 93

RESULT 5
E90116
mRNA capping enzyme [imported] - Guillardia theta nucleomorph

C/Species: nucleomorph Guillardia theta

A/Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001

C/Accession: E90116

R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re

Nature 410, 1091-1096, 2001

A/Title: The highly reduced genome of an enslaved algal nucleus.

A/Reference number: A99082; MUID:11323671; PMID:11323671

A/Accession: E90116

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-357 <DOU>

A/Cross-references: GB:AJ010592; NID:g12580792; PIDN:CAC27110.1; GSPDB:GN00151

C/Genetics:
A/Map position: 2

A/Genome: nucleomorph

C/Keywords: nucleomorph

Query Match 53.0%; Score 44; DB 2; Length 357;
Best Local Similarity 57.1%; Pred. No. 5.6;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTFD 14
|||
Db 91 KMPMYIAGYKTFD 104

RESULT 6
A70384
Glutamate N-acetyltransferase - Aquifex aeolicus

C/Species: Aquifex aeolicus

C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999

C/Accession: A70384

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

Nature 392, 353-358, 1998

A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A/Reference number: A70300; MUID:98196666; PMID:9537320

A/Accession: A70384

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-379 <AOB>

A/Cross-references: GB:AE000716; NID:g2983478; PIDN:AA07058.1; PID:g2983479; GB:AE0006

A/Experimental source: strain VFS

C/Genetics:
A/Genome: argu

C/Superfamily: glutamate N-acetyltransferase argu

Query Match 53.0%; Score 44; DB 2; Length 379;
Best Local Similarity 40.0%; Pred. No. 6;

Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTFDGR 15
|||
Db 313 KLEIYGVGLYLVK 327

RESULT 7
S76571
Hypothetical protein - Synechocystis sp. (strain PCC 6803)

C/Species: Synechocystis sp.

A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C/Accession: S76571
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76571
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <KAN>
A:Cross-references: EMBL:D64002; GB:AB001339; NID:G1001612; PIDN:BA10417.1; PID:dt01106
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match
Best Local Similarity 51.8%; Score 43; DB 2; Length 419;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 GYKTFDGR 15
||:|||||
Db 183 GYKTFDGR 190

RESULT 8

S26863

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Podospora anserina

C:Species: Podospora anserina

C>Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 03-Jun-2002

C:Accession: S26863

R:Ridder, R.; Osiewicz, H.D.

Curr. Genet. 21, 207-213, 1992

A:Title: Sequence analysis of the gene coding for glyceraldehyde-3-phosphate dehydrogenase

A:Reference number: S26863; MUID:92224295; PMID:1563046

A:Accession: S26863

A:Molecule type: DNA

A:Residues: 1-337 <RID>

A:Cross-references: EMBL:X62824; NID:G3101; PIDN:CAA44635.1; PID:G3102

C:Genetics:

A:Gene: gpd

A:Insertions: 42/3

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase

Query Match
Best Local Similarity 50.6%; Score 42; DB 2; Length 337;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 3 PMYIAGY-KTFDGR 15
||:|||||
Db 128 PMYIAGYKTFDGR 142

RESULT 9

S40610

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - ergot fungus

C:Species: Claviceps purpurea (ergot fungus)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 03-Jun-2002

C:Accession: S40610

R:Jungeheusing, U.; Arntz, C.; Smit, R.; Tudzynski, P.

Curr. Genet. 25, 101-106, 1994

A:Title: The Claviceps purpurea glyceraldehyde-3-phosphate dehydrogenase gene: cloning,

A:Reference number: S40610; MUID:94373854; PMID:8087877

A:Accession: S40610

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-337 <JUN>

A:Cross-references: EMBL:X73282; NID:G452310; PIDN:CAA51721.1; PID:G452311

C:Genetics:

A:Insertions: 42/3

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: gluconeogenesis; glycolysis; oxidoreductase

Query Match
Best Local Similarity 50.6%; Score 42; DB 2; Length 337;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 3 PMYIAGY-KTFDGR 15
||:|||||
Db 128 PMYIAGYKTFDGR 142

RESULT 10

T30889

valine-tRNA ligase (EC 6.1.1.9) - Giardia lamblia

C:Species: Giardia lamblia

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 03-Jun-2002

C:Accession: T30889

R:Hashimoto, T.; Sanchez, J.B.; Shirakura, T.; Muller, M.; Haegawa, M.

Proc. Natl. Acad. Sci. U.S.A. 95, 6860-6865, 1998

A:Title: Secondary absence of mitochondria in Giardia lamblia and Trichomonas vaginalis

A:Reference number: Z20923; MUID:96884022; PMID:9618503

A:Accession: T30889

A:Molecule type: DNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-1218 <HAS>

A:Cross-references: EMBL:AB008525; PIDN:BA28839.1

C:Superfamily: valine-tRNA ligase

C:Keywords: ligase; protein biosynthesis

Query Match
Best Local Similarity 50.6%; Score 42; DB 2; Length 1218;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 PMYIAGYKTFDGR 15
||:|||||
Db 465 LPARFAGYKTFDGR 478

RESULT 11

T44562

catalase (EC 1.11.1.6) HPI [imported] - Halobacterium salinarum

N:Alternate names: hydroperoxidase

C:Species: Halobacterium salinarum

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T44562

R:Salin, M.L.; Long, S.

submitted to the EMBL Data Library, June 1998

A:Description: Hydroperoxidase.

A:Reference number: Z22793

A:Accession: T44562

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-720 <SAL>

A:Cross-references: EMBL:AF069761; PIDN:AA233534.1

C:Superfamily: catalase HPI

C:Keywords: heme; iron; metalloprotein; oxidoreductase

F/83/Active site: His (distal axial ligand) #status predicted

F/248/Binding site: heme iron (His) (proximal axial ligand) #status predicted

F/300/Active site: Trp (tryptophyl radical intermediate) #status predicted

Query Match
Best Local Similarity 50.0%; Score 41.5; DB 2; Length 720;
Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 KMPYIAGY-KTFDGR 15
||:|||||
Db 79 RYAMHSAGYKTFDGR 94

RESULT 12

S49369

mobilization protein - Campylobacter coli plasmid pCCT1 and pCCT2

C:Species: Campylobacter coli

C>Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 08-Oct-1999

C:Accession: S49369; S49372

R:Stonnet, V.; Quesdon, J.L.

submitted to the EMBL Data Library, October 1994

A:Description: Nucleotide sequences and molecular organization of the two small cryptic

A:Reference number: S49369

A:Accession: S49369

A:Status: preliminary

A:Molecule type: DNA

A/Residues: 1-321 <STO>
 A/Cross-references: EMBL:X82080; NID:G557224; PIDN:CAA57597.1; PID:G557227
 A/Experimental source: plasmid pCCT2
 A/Genetics: G1
 A/Accession: S49372
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-321 <STN>
 A/Cross-references: EMBL:X82079; NID:G557228; PIDN:CAA57594.1; PID:G557231
 A/Experimental source: plasmid pCCT1
 A/Genetics: G2
 A/Genetics: <G1>
 A/Gene: mob
 A/Genome: plasmid pCCT2
 C/Genetics: <G2>
 A/Gene: mob
 A/Genome: plasmid pCCT1
 C/Keywords: plasmid mobilization

Query Match 49.4%; Score 41; DB 2; Length 321;
 Best Local Similarity 58.3%; Pred. No. 17;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTF 12
 Db 94 KOPYIODYKTY 105

RESULT 13
 T06728
 pectate lyase (EC 4.2.2.2) F28P10.100 - Arabidopsis thaliana
 N/Alternate names: protein F28P10.100
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
 C/Accession: T06728
 R/Querier, F.; Choisme, N.; Robert, C.; Broctier, P.; Wincker, P.; Catolico, L.; Artigues
 submitted to the Protein Sequence Database, April 1999
 A/Reference number: Z15793
 A/Accession: T06728
 A/Molecule type: DNA
 A/Residues: 1-542 <QUB>
 A/Cross-references: EMBL:AL049655
 A/Experimental source: cultivar Columbia; BAC clone F28P10
 C/Genetics:
 A/Map position: 3
 A/Introns: 46/2; 346/3; 413/2; 480/2
 A/Note: F28P10.100
 C/Superfamily: pectate lyase LAT59
 C/Keywords: carbon-oxygen lyase

Query Match 49.4%; Score 41; DB 2; Length 542;
 Best Local Similarity 53.3%; Pred. No. 30;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTFDGR 15
 Db 160 KOELIINSYKTLDDGR 174

RESULT 14
 S38653
 transposase - Klebsiella pneumoniae
 C/Species: Klebsiella pneumoniae
 C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999
 C/Accession: S38653
 R/Rinkel, M.
 submitted to the EMBL Data Library, November 1993
 A/Reference number: S38652
 A/Accession: S38653
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-952 <RIN>
 A/Cross-references: EMBL:X75779; NID:G415974; PIDN:CAA53399.1; PID:G415976

C/Superfamily: transposase Tn21
 Query Match 49.4%; Score 41; DB 2; Length 952;
 Best Local Similarity 80.0%; Pred. No. 54;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 6 IAGYKTFDGR 15
 Db 110 IYGYKTFDGR 119

RESULT 15
 T30881
 dynein heavy chain isotype 3A - sea urchin (Tripneustes gratilla) (fragment)
 N/Contains: dynein ATPase (EC 3.6.4.2)
 C/Species: Tripneustes gratilla
 C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 19-Apr-2002
 C/Accession: T30881
 R/Gibbons, B.H.; Asai, D.U.; Tang, W.J.; Hays, T.S.; Gibbons, I.R.
 Mol. Biol. Cell 5, 57-70, 1994
 A/Title: Phylogeny and expression of axonemal and cytoplasmic dynein genes in sea urchin
 A/Reference number: Z20814; M01D:94243035; PMID:8186465
 A/Accession: T30881
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1071 <GIB>
 A/Cross-references: EMBL:U03970; NID:G433368; PID:G433369; PIDN:AAA63584.1
 A/Experimental source: tissue type embryo; developmental stage mesenchyme blastula
 C/Genetics:
 A/Gene: DYH3A
 C/Superfamily: dynein heavy chain, ciliary
 C/Keywords: ATP; hydrolase; microtubule binding; P-loop

Query Match 49.4%; Score 41; DB 2; Length 1071;
 Best Local Similarity 58.3%; Pred. No. 61;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTF 12
 Db 1021 KLASFIAGYKSF 1032

Search completed: April 20, 2003, 13:15:44
 Job time : 8.07895 secs

POLLEN, THE MOST COMMON POLLEN ALLERGEN IN JAPAN.

B DIFFERS IN SIX POSITIONS.
 CC ENTITIES BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D26544; BAA05542.1; -
 CC EMBL: D26545; BAA05543.1; -
 CC EMBL: D34639; BAA07020.1; -
 CC PIR: A44773; A44773.
 CC DR GLYCOSULEDB; P18632; -
 CC DR InterPro; IPR002022; Amb_allergen.
 CC DR Pfam; PF00544; pec_lyase; 1.
 CC DR PRINTS; PR00807; AMBALLERGEN.
 CC Allergen; Glycoprotein; Multigene family; Signal.
 KW SIGNAL
 FT SIGNAL 1 21
 FT CHAIN 22 374
 FT CARBOHYD 158 158
 FT CARBOHYD 191 191
 FT CARBOHYD 293 293
 FT CARBOHYD 354 354
 FT CARBOHYD 361 361
 FT VARIANT 12 12
 FT VARIANT 143 143
 FT VARIANT 202 202
 FT VARIANT 221 221
 FT VARIANT 358 358
 FT VARIANT 361 361
 FT SEQUENCE 374 AA; 40645 MW; 74AB25950248F56F CRC64;
 SQ
 Query Match 100.0%; Score 83; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KMPMYIAGYKTFDGR 15
 DB 102 KMPMYIAGYKTFDGR 116
 RESULT 2
 MPAL_CUPAR STANDARD; PRT; 346 AA.
 AC Q9SCG9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Cup a 1.
 OS Cupressus arizonica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=99011;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20571526; Pubmed=11122214;
 RA Acetuno E., Del Pozo V., Minguet A., Arieta I., Cortegano I.,
 RA Cardaba B., Gallardo S., Rojo M., Palomino P., Lahoz C.;
 RT "Molecular cloning of major allergen from Cupressus arizonica pollen:
 RT Cup a 1.";
 RL Clin. Exp. Allergy 30:1750-1758(2000).
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ243570; CAB62551.1; -
 CC InterPro; IPR002022; Amb_allergen.
 CC Pfam; PF00544; pec_lyase; 1.
 CC DR PRINTS; PR00807; AMBALLERGEN.
 CC Allergen; Glycoprotein.
 KW CARBOHYD 127 127
 FT CARBOHYD 157 157
 FT CARBOHYD 272 272
 FT SEQUENCE 346 AA; 37589 MW; F1281DCDA1D5DFD0 CRC64;
 SQ
 Query Match 83.1%; Score 69; DB 1; Length 346;
 Best Local Similarity 73.3%; Pred. No. 0.00011;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 KMPMYIAGYKTFDGR 15
 DB 81 QMPMYIAGYKTFDGR 95
 RESULT 3
 MPAL_JUNAS STANDARD; PRT; 367 AA.
 AC P81234; Q9ZNU7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Jun a 1 precursor.
 OS Juniperus ashei (Ozark white cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxID=1101;
 RN (1)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-38; 42-50; 58-80; 88-94;
 RP 117-124; 134-140; 160-164; 256-263 AND 322-325;
 RX MEDLINE=99414163; Pubmed=10482836;
 RA Midoro-Horinuti T.M., Goldblum R.M., Kurosky A., Wood T.G.,
 RA Brooks E.G.;
 RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 RT allergen, Jun a 1.";
 RL J. Allergy Clin. Immunol. 104:613-617(1999).
 RN (2)
 RP SEQUENCE OF 22-50.
 RC TISSUE-Pollen;
 RX MEDLINE=99414162; Pubmed=10482835;
 RA Midoro-Horinuti T., Goldblum R.M., Kurosky A., Goetz D.W.,
 RA Brooks E.G.;
 RT "Isolation and characterization of the mountain cedar (Juniperus
 RT ashei) pollen major allergen, Jun a 1.";
 RL J. Allergy Clin. Immunol. 104:608-612(1999).
 CC -1- DISBASE: THIS PROTEIN CAUSES SEVERAL SEASONAL ALLERGIC RHINITIS
 CC IN NORTH AMERICA.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF106662; AAD03609.1; -
 CC EMBL: AF106662; AAD03608.1; -
 CC InterPro; IPR002022; Amb_allergen.
 CC Pfam; PF00544; pec_lyase; 1.
 CC DR PRINTS; PR00807; AMBALLERGEN.
 KW Allergen; Glycoprotein; Signal.

FT SIGNAL 1 21
 FT CHAIN 22 367 MAJOR POLLEN ALLERGEN JUN A 1.
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 367 AA; 39824 MW; FC9881E75662849 CRC64;

Query Match
 Best Local Similarity 81.9%; Score 68; DB 1; Length 367;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMPYIAGYKTFDGR 15
 DB 102 KMPYIAGYKTFDGR 116

RESULT 4
 MEAL CHAQB STANDARD; PRT; 375 AA.

ID ID_MPAI CHAQB STANDARD; PRT; 375 AA.
 AC 096385;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major pollen allergen Cha o 1 precursor
 OS Chamaecyparis obtusa (Japanese cypress).
 OC Eukaryota; Viridiplantae; Streptophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
 OC Chamaecyparis;
 CC NCBI_TaxID=13415;
 CC [1]
 CC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC TISSUE=Pollen;
 RX MEDLINE=96265194; PubMed=8676896;
 RA Suzuki M., Komiyama N., Itoh H., Sone T., Kuno K., Takagi I.,
 RA Ohta N.;

RT "Purification, characterization and molecular cloning of Cha o 1, a
 RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen."
 RT Mol. Immunol. 33:451-460 (1996).
 CC -1 SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL, D45404; BAA08246.1; -
 CC InterPro: IPR002022; Amb allergen.
 CC Pfam: PF00544; pec lyase; 1.
 CC PRINTS: PRO0807; AMBALLERGEN.
 CC Allergen; Glycoprotein; Signal.

FT CHAIN 1 21
 FT SIGNAL 22 375 MAJOR POLLEN ALLERGEN CHA O 1.
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF7066DBF CRC64;

Query Match
 Best Local Similarity 72.3%; Score 60; DB 1; Length 375;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 MPMYIAGYKTFDGR 15
 DB 103 MPMYIAGYKTFDGR 116

RESULT 5
 GLGA_SYNY3

ID GLGA_SYNY3 STANDARD; PRT; 477 AA.

AC P74521;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glycogen synthase (EC 2.4.1.21) (Starch [bacterial glycogen]
 DE synthase).
 GN GLGA OR SLU0945.
 OS Synchocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
 CC NCBI_TaxID=1148;
 CC [1]
 CC SEQUENCE FROM N.A.

CC MEDLINE=97061201; PubMed=8905231;
 CC RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 CC RA Miyajima N., Hiroseawa M., Sugitani M., Sasamoto S., Kimura T.,
 CC RA Hosouchi T., Matsuno A., Muraki C., Nakazaki N., Nario K.,
 CC RA Okumura S., Shimo S., Takeuchi C., Wada T., Matanabe A.,
 CC RA Yamada M., Yasuda M., Tabata S.;
 CC RT "Sequence analysis of the genome of the unicellular cyanobacterium
 CC Synchocystis sp. strain PCC6803. II. Sequence determination of the
 CC RT entire genome and assignment of potential protein-coding regions."
 CC RL DNA Res. 3:109-136 (1996).
 CC -1 FUNCTION: Synthesizes alpha-1,4-glucan chains using ADP-glucose.
 CC -1 CATALYTIC ACTIVITY: ADP-glucose + ((1,4)-alpha-D-glucosyl) (N) =
 CC ADP + ((1,4)-alpha-D-glucosyl) (N+1).
 CC -1 PATHWAY: Glycogen biosynthesis; second step.
 CC -1 SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL, D90915; BAA18625.1; -
 CC InterPro: IPR001296; Glycosyltransferase.
 CC Pfam, PF00534; Glycosyltransferase; 1.
 CC KX Glycogen biosynthesis; Transferase; Glycosyltransferase;
 CC Complete proteome.
 CC FT BINDING 15 15 ADP-GLUCOSE (BY SIMILARITY).
 SQ SEQUENCE 477 AA; 54346 MW; 40C1B6013938E032 CRC64;

Query Match
 Best Local Similarity 55.4%; Score 46; DB 1; Length 477;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 KMPYIAGYKTFDGR 15
 DB 79 KIPYIAGYKTFDGR 93

RESULT 6
 G3P CLAPU STANDARD; PRT; 337 AA.
 AC 000584;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GPD-1.
 OS Claviceps purpurea (Brgot fungus).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocerales; Clavicipitaceae; Claviceps.
 CC NCBI_TaxID=5111;
 CC [1]
 CC SEQUENCE FROM N.A.

CC STRAIN=75;
 CC MEDLINE=94373854; PubMed=8087877;
 CC Jungheunelung U., Arntz C., Smit R., Tudzynski P.;


```

RT      "the Claviceps purpurea glyceraldehyde-3-phosphate dehydrogenase
RT      gene: cloning, characterization, and use for the improvement of a
RT      dominant selection system.";
RL      Curr. Genet. 25:101-106 (1994).
CC      -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC      NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC      -1- PATHWAY: Second phase of glycolysis; first step.
CC      -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC      DEHYDROGENASE FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL: X73282; CA51721.1; -.
DR      HSRP; P56649; IDSS.
DR      InterPro; IPR00173; GAP_dhydrogenase.
DR      Pfam; PF00044; gpdh; 1.
DR      Pfam; PF02800; gpdh_C; 1.
DR      PRINTS; PR0078; G3PDHDEGNASE.
DR      PROSITE; PS00071; GAPDH; 1.
KW      GLYCOLYSIS; Oxidoreductase; NAD.
FT      BINDING 151 151
FT      ACT_SITE 178 178
FT      ACTIVES_THIOL_GROUP_DURING_CATALYSIS.
SQ      SEQUENCE 337 AA; 36029 MW; 1C2360B9701E26B9 CRC64;

Query Match 50.6%; Score 42; DB 1; Length 337;
Best local Similarity 53.3%; Pred. No. 5.6;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY      3 PMYIAGY--KTFDGR 15
DB      128 PMYVMGVNERTYDGK 142
      |||: |||: |||:
      |||: |||: |||:

RESULT 7
G3P_PODAN STANDARD; PRT; 337 AA.
ID      G3P_PODAN
DC      P32637;
DT      01-OCT-1993 (Rel. 27, Created)
DT      01-OCT-1993 (Rel. 27, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN      GPD.
OS      Podospora anserina.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC      Sordariales; Lasiosphaeriaceae; Podospora.
OX      NCBI TaxId=5145;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-92224295; PubMed=1563046;
RA      Rüdger R., Osiewacz H.D.;
RT      "Sequence analysis of the gene coding for glyceraldehyde-3-phosphate
RT      dehydrogenase (gpd) of Podospora anserina: use of homologous
RT      regulatory sequences to improve transformation efficiency.";
RL      Curr. Genet. 21:207-213 (1992).
CC      -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC      NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC      -1- PATHWAY: Second phase of glycolysis; first step.
CC      -1- SUBUNIT: HOMOTETRAMER.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC      DEHYDROGENASE FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its

```

```

CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X62824; CAAA635.1;  -
DR      PIR; S26863; S26863.
DR      HSSP; P56649; IDS5.
DR      InterPro; IPR000173; GAP_dhdrogenase.
DR      Pfam; PF00044; gpdh, 1.
DR      Pfam; PF02800; gpdh.C; 1.
DR      PRINTS; PR00078; GAPDHGNASE.
DR      PROSITE; PS00071; GAPDH; 1.
DR      GLYCOLYSIS; Oxidoreductase; NAD.
KW      GLYCERALDEHYDE 3-PHOSPHATE.
FT      BINDING 151 151
FT      ACT SITE 178 178
FT      ACT SITE 36211 36211 MW; C9A0FIDCC05A643B CRG64;
SQ
SQ
Query Match          50.6%;      Score 42; DB 1; Length 337;
Best Local Similarity 53.3%;      Pred. NO. 5.6;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 1
QY      3  PMYIAGY--KTFDGR 15
      |||::|||::|||:
Db      128 PMYIVMGVNEKTYDGK 142

```

RESULT 8
CATA_HALNT1 STANDARD; PRT; 720 AA.
ID CATA_HALNT1
AC 073955; 09HH95;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peroxidase/catalase (EC 1.11.1.6) (Catalase-peroxidase)
DE (Hydroperoxidase).
DE PERA OR VNG6294G.
OS Halobacterium sp. (strain NRC-1), and
OS Halobacterium salinarum.
OC Plasmid pNRC200
OC Archaea, Euryarchaeota, Halobacteriia, Halobacteriales;
OC Halobacteriaceae, Halobacterium.
OX NCBI_TaxId=64091, 2242;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=NRC-1; PLASMID=pNRC200;
RC MEDLINE=20504483; PubMed=110165950;
RX NG W.V., Kennedy S.P., Mahairs G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasry S.R., Baliga N.S., Thorsson V., Shroya J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Denson M.J., Hough D.W.,
RA Meddoles D.G., Jablonski P.E., Krebs W.P., Amegvane C.M., Dale H.,
RA Iembaerger T.A., Peck R.F., Danielshroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S.F., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RA genome sequence of Halobacterium species NRC-1.";
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
[2]
RN SEQUENCE FROM N.A.
RP SPECIES=H.salinarum;
RC Salin M.L., Long S.;
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RL FUNCTION: BIFUNCTIONAL, EXHIBITING BOTH A CATALASE AND
CC BROAD-SPECTRUM PEROXIDASE ACTIVITIES (BY SIMILARITY).
CC
CC -1 CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC
CC -1 COFACTOR: Binds 2 prothème IX and 2 iron ions per tetramer (By
CC similarity).
CC
CC -1 SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. BACTERIAL
CC PEROXIDASE/CATALASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE005159; AAC20931.1; -
EMBL; AF069761; AAC23534.1; -
HSSP; P00431; IRYC.
InterPro; IPR000763; Bac_citase/prase.
InterPro; IPR002016; Peroxidase.
Pfam; PF00141; peroxidase; 1.
PRINTS; P00458; PEROXIDASE.
TIGRfam; TIGR00198; cat_per_HPI; 1.
PROSITE; PS00435; PEROXIDASE_1; 1.
PROSITE; PS00436; PEROXIDASE_2; 1.
Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide; Plasmid;
Complete proteome.
ACT SITE 79 79 BY SIMILARITY.
ACT SITE 83 83 BY SIMILARITY.
METAL 248 248 IRON (PROTOME IX AXIAL LIGAND).
SEQUENCE 720 AA; 80476 MW; FB88823BCD3CB2F9 CRC64;
Query Match 50.0%; Score 41.5; DB 1; Length 720;
Best Local Similarity 56.2%; Pred. No. 14;
Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;
OY 1 KMPWYIAG-YKTFDGR 15
DB 79 RWAMSHAGTYRTFDGR 94

RESULT 9
G3P_CRYPA STANDARD; PRT; 337 AA.
ID G3P_CRYPA
AC P19089;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GPD-1).
OS Cyphonectria parasitica (Cheesnut blight fungus) (Endothia parasitica).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Diaporthales; Valsaceae; Cryphonectria.
OX NCBI_TaxID=5116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91016863; PubMed=2216743;
RA Choi G.H., Nuss D.L.;
RT "Nucleotide sequence of the glyceraldehyde-3-phosphate dehydrogenase gene from *Cyphonectria parasitica*.";
RL Nucleic Acids Res. 18:5566-5566(1990).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X53996; CAA37943.1; -
PIR; S11447; DEJUC.
HSSP; P56649; IDS.
InterPro; IPR000173; GAP_dhhydrogenase.
Pfam; PF00044; spdh; 1.
Pfam; PF02800; spdh_C; 1.

PRINTS; PRO0078; G3PDHRCNASE.
PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD.
FT BINDING 151 151
FT ACT SITE 178 178
SEQUENCE 337 AA; 36194 MW; 38C04A557E68E6 CRC64;
Query Match 48.2%; Score 40; DB 1; Length 337;
Best Local Similarity 57.1%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;
OY 3 PMVIAGY--KTFFG 14
DB 128 PMVWGVNERTYDG 141

RESULT 10
HYBE_ECOLI STANDARD; PRT; 162 AA.
ID HYBE_ECOLI
AC P37183;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hydrogenase-2 operon protein hybe.
GN HYBE OR B2992.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / TGI;
RC MEDLINE=94292472; PubMed=8021226;
RA Meron N.K., Chateaus C.Y., Dervartanian M., Wendt J.C.,
RA Shanmugam K.T., Peck H.D. Jr., Przybyla A.E.;
RT "Cloning, sequencing, and mutational analysis of the hyb operon encoding *Escherichia coli* hydrogenase 2.";
RL J. Bacteriol. 176:4416-4423(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RL Science 277:1234-1238(1997).
CC -1- SIMILARITY: BELONGS TO THE HUPV FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U09177; AAA21593.1; -
EMBL; U28377; AAA69159.1; -
EMBL; AE000382; AAC76028.1; -
Ecogene; E011803; hybe.
KW Complete proteome.
FT CONFLICT 7 7 A -> G (IN REF. 2).
SEQUENCE 162 AA; 17977 MW; CFPD10617D815570 CRC64;
Query Match 47.0%; Score 39; DB 1; Length 162;
Best Local Similarity 35.7%; Pred. No. 9.3;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
OY 2 MPMYINGYKTFDGR 15
DB 38 MPMYVSDFTLFEQ 51

RC STRAIN=CV. VF36; TISSUE=distill;
 RX MEDLINE=91117185; PubMed=2277637;
 RA Budelier K.A., Smith A.G., Gasser C.S.;
 RT "Regulation of a stylar transmitting tissue-specific gene in
 RL wild-type and transgenic tomato and tobacco.";
 CC Mol. Gen. Genet. 224:183-192(1990).
 CC -1- FUNCTION: MAY HAVE A ROLE IN THE DEVELOPMENT OF THE TRANSMITTING
 CC TISSUE OF THE STYLE AND/OR IN THE EVENTS RELATED TO POLLINATION
 CC SUCH AS SOME ASPECT IN THE FACILITATION OF COMPATIBLE POLLEN TUBE
 CC GROWTH.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN THE PISTIL WHERE IT IS
 CC FOUND IN THE OUTER FIVE LAYERS OF THE STRANDS OF TRANSMITTING
 CC TISSUE WITHIN THE UPPER TWO-THIRDS OF THE STYLE. FOUND AT MUCH
 CC LOWER LEVELS IN THE ANthers AND VEGETATIVE STEMS.
 CC -1- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND DURING ANTHESIS.
 CC -1- SIMILARITY: 54% IDENTICAL TO TOMATO PROTEINS P59 (AC P15722)
 CC AND P56 (AC P15721).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X55193; CA38979.1; -.
 DR F1R; S12209; S12209.
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase_1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KM Signal.
 FT SIGNAL. 1 20
 FT CHAIN 1 404
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 404 AA; 44298 MW; B26E69B128D8675 CRC64;
 QY 1 KMPMYIAGYKTFDGR 15
 Db 134 KQELVWMSYKTIIDGR 148
 RESULT 14
 IDHP SCHPO STANDARD; PRT; 418 AA.
 AC 014254;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable isocitrate dehydrogenase [NADP], mitochondrial precursor
 DE (EC 1.1.1.42) (Oxalosuccinate decarboxylase) (IDH) (NADP+-specific
 DE ICDH) (IDP).
 GN SPAC6310.08.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OK NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou K., Beat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks J., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckie E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leacher S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkhardt G., Aert R., Robben J., Grymoprez B.,
 RA Wellens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel K., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Mandut B., Purnelle B.,
 RA Galibert F., Aves S.J., Xiang Z., Gloux S., Lelaune V., Motlier S.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Folsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Pallesen I., Potashkin J.,
 RA Shpakovski G.V., Useery D., Barrett B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 CC Nature 415:871-880(2002).
 CC -1- CATALYTIC ACTIVITY: Isocitrate + NADP(+) = 2-oxoglutarate + CO(2)
 CC + NADPH.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE
 CC DEHYDROGENASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z98603; CAB11294.1; -.
 DR InterPro; IPR001804; Isodh.
 DR InterPro; IPR004730; NADp_idh_euk.
 DR Pfam; PF00180; isodh_1.
 DR PRODOM; PD003173; NADp_idh_euk; 1.
 DR TIGRPFAM; TIGR00127; nadp_idh_euk; 1.
 DR PROSITE; PS00470; IDH IMDH; 1.
 KM Oxidoreductase; NADP; Glyoxylate lyase; Tricarboxylic acid cycle;
 KW Transil peptide; Mitochondrion.
 FT TRANSIT 1 2
 FT CHAIN 1 418
 FT ACT SITE 104 104 PROBABLE ISOCITRATE DEHYDROGENASE [NADP].
 FT SEQUENCE 418 AA; 47293 MW; 6E114B379C8AD61E CRC64;
 QY 1 KMPMYIAGYKTFDGR 15
 Db 214 KQELVWMSYKTIIDGR 232
 RESULT 15
 G6PB ORYZA STANDARD; PRT; 568 AA.
 AC P42863;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glucose-6-phosphate isomerase, cytosolic B (EC 5.3.1.9) (GPI-B)
 DE (Phosphoglucose isomerase B) (PGI-B) (Phosphohexose isomerase B) (PHI-
 DE B).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OK NCBI_TaxID=4530;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96293857; PubMed=8722567;
RA Nozue F., Umeda M., Nagamura Y., Minobe Y., Uchimiya H.;
RT "Characterization of cDNA encoding for phosphoglucose isomerase of
RT rice (Oryza sativa L.).";
RT DNA Seq. 6:127-135(1996).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GPI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D45218; BAA08149.1; -.
DR HSP; Q5N1E2; IHOX.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
DR PROSITE; PS00174; P_GLUCOSE_ISOMERASE_2; 1.
KW Gluconeogenesis; Glycolysis; BY SIMILARITY.
FT ACT_SITE 392
FT ACT_SITE 517
FT ACT_SITE 517 BY SIMILARITY.
SQ SEQUENCE 568 AA; 62419 MW; C84C32EF265CA12 CRC64;

```

```

Query Match 45.8%; Score 38; DB 1; Length 568;
Best Local Similarity 42.9%; Pred. No. 46;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 1 KMPYIAGYKTPDG 14
|:|::|||
Db 455 KVPBHLISKTFQG 468

```

Search completed: April 20, 2003, 13:07:34
Job time : 3.92105 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 12.6711 Seconds
(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524D-31

Perfect score: 83

Sequence: 1 KMPMYIAGYKTFDGR 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP tvirus:*
16: SP bacteriophage:*
17: SP archaeop:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	100.0	374	10	Q8RUR1 cryptomeria
2	69	83.1	367	10	Q93XL6 cupressus a
3	68	81.9	367	10	Q91LT2 juniperus v
4	68	81.9	367	10	Q91LT1 juniperus v
5	68	81.9	367	10	Q9M4S6 cupressus s
6	68	81.9	367	10	Q9M4S3 cupressus s
7	68	81.9	367	10	Q9M4S2 cupressus s
8	68	81.9	367	10	Q9M4S1 juniperus o
9	68	81.9	367	10	Q93XL5 juniperus o
10	64	77.1	367	10	Q9M4S5 cupressus s
11	48	57.8	157	16	Q9A413 canlobacter
12	46	55.4	392	10	Q9FM66 arabisopsis
13	45	54.2	122	16	Q8RE02 fusobacteri
14	45	54.2	3508	5	Q9VH97 diosophila
15	44	53.0	357	10	Q9AVW1 guillardia
16	44	53.0	379	16	O67100 aquifex aeo

17	43	51.8	419	16	Q55770 synechocyst
18	42	50.6	690	13	Q9PVW5 oryzias lat
19	42	50.6	1218	5	O77443 giardia lam
20	41	49.4	181	10	Q9ZUY5 arabidopsis
21	41	49.4	203	10	Q9M4T3 arabidopsis
22	41	49.4	206	16	Q9EWV6 streptomyce
23	41	49.4	273	12	Q91C42 arabidopsis
24	41	49.4	273	12	Q91GAI1 arabidopsis
25	41	49.4	321	2	O57116 campylobact
26	41	49.4	347	16	Q8XYK8 arabidopsis
27	41	49.4	419	16	Q8XYK8 arabidopsis
28	41	49.4	501	10	Q93Z04 arabidopsis
29	41	49.4	542	10	Q9SV40 arabidopsis
30	41	49.4	952	2	Q48490 klebsiella
31	41	49.4	1071	5	Q27803 tripeustes
32	40	48.2	288	17	Q8U2S7 arabidopsis
33	40	48.2	336	16	Q8RFP1 fusobacteri
34	40	48.2	353	4	Q96OH5 homo sapien
35	40	48.2	353	4	Q8WY02 homo sapien
36	40	48.2	353	4	Q8WY02 homo sapien
37	40	48.2	368	10	Q9C8G4 homo sapien
38	40	48.2	378	4	Q9C0J1 homo sapien
39	40	48.2	477	17	Q8TW11 homo sapien
40	40	48.2	974	17	Q981D6 methanosarc
41	39	47.0	162	16	Q8XBW0 escherichia
42	39	47.0	289	12	Q9Y1D4 schizosacch
43	39	47.0	410	3	Q9URV7 emericella
44	39	47.0	419	3	Q9HEV1 rhizobium m
45	39	47.0	445	16	Q92T77 rhizobium m

ALIGNMENTS

RESULT 1
ID Q8RUR1 PRELIMINARY; PRT; 374 AA.
AC Q8RUR1;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Cry j 1 precursor.
GN Cry j 1.1 OR Cry j 1.2.
OS Cryptomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OX NCBI_TaxID=3369;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=POLLEN;
RA Futamura N., Shinohara K.;
RT "Isolation and characterization of cDNAs encoding major allergen Cry j
1 from Cryptomeria japonica pollen.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB081309; BAB86286.1; -;
DR EMBL; AB081310; BAB86287.1; -;
KW Signal.
FT SIGNAL. 1 21 POTENTIAL.
FT CHAIN 22 374 CRY J 1.
SQ SEQUENCE 374 AA, 40720 MW, 9000085D24BF2BD4 CRC64;

Query Match 100.0%; Score 83; DB 10; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMPMYIAGYKTFDGR 15
Db 102 KMPMYIAGYKTFDGR 116

RESULT 2
Q93XL6 PRELIMINARY; PRT; 367 AA.

```

AC 093XL6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative allergen Cup a 1 precursor.
GN CUP A 1.
OS Cupressus arizonica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OK NCBI_TaxID=49011;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=POLLEN;
RA Butterworth C., Di Felice G., Pini C.;
RT "Cloning of Cupressus arizonica major allergen.";
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ278498; CAC37790.2; -.
FT SIGNAL.
FT CHAIN 22 367 PUTATIVE ALLERGEN CUP A 1.
SQ SEQUENCE 367 AA; 39809 MW; AEF97260423A9F28 CRC64;

Query Match 83.1%; Score 69; DB 10; Length 367;
Best Local Similarity 73.3%; Pred. No. 0.00087;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KMPMTIAGYKTFDGR 15
ID 102 KMPLYVAGHKITDGR 116
DB 102 KMPLYVAGHKITDGR 116

RESULT 3
O9LIT2 PRELIMINARY; PRT; 367 AA.
AC O9LIT2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-2.
DS Juniperus virginiana (Eastern red cedar).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OK NCBI_TaxID=39584;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=21315424; PubMed=11422137;
RA Midoro-Horicuti T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RT Clin. Exp. Allergy 31:771-778(2001).
RL EMBL; AF151427; AA80164.1; -.
DR EMBL; AF151427; AA80164.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF89586 CRC64;

Query Match 81.9%; Score 68; DB 10; Length 367;
Best Local Similarity 73.3%; Pred. No. 0.0013;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KMPMTIAGYKTFDGR 15
ID 102 KMPLYVAGHKITDGR 116
DB 102 KMPLYVAGHKITDGR 116

RESULT 4
O9LIT1 PRELIMINARY; PRT; 367 AA.
AC O9LIT1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-1.

```

```

OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OK NCBI_TaxID=39584;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=21315424; PubMed=11422137;
RA Midoro-Horicuti T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RT Clin. Exp. Allergy 31:771-778(2001).
RL EMBL; AF151429; AA80166.1; -.
DR EMBL; AF151429; AA80166.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;

Query Match 81.9%; Score 68; DB 10; Length 367;
Best Local Similarity 73.3%; Pred. No. 0.0013;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KMPMTIAGYKTFDGR 15
ID 102 KMPLYVAGHKITDGR 116
DB 102 KMPLYVAGHKITDGR 116

RESULT 5
O9M4S6 PRELIMINARY; PRT; 367 AA.
ID O9M4S6;
AC O9M4S6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
DS Cupressus sempervirens.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OK NCBI_TaxID=13469;
RN (1)
RP SEQUENCE FROM N.A.
RC Monstave R.I., Villalba M., Rodriguez R.;
RA "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF257491; AAF72625.1; -.
DR EMBL; AF257491; AAF72625.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39835 MW; BA89C60108C2C5A3 CRC64;

Query Match 81.9%; Score 68; DB 10; Length 367;
Best Local Similarity 73.3%; Pred. No. 0.0013;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KMPMTIAGYKTFDGR 15
ID 102 KMPLYVAGHKITDGR 116
DB 102 KMPLYVAGHKITDGR 116

RESULT 6
O9M4S4 PRELIMINARY; PRT; 367 AA.
ID O9M4S4;
AC O9M4S4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
DS Cupressus sempervirens.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OK NCBI_TaxID=13469;

```

RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257493; AAF72627.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SO SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 81.9%; Score 68; DB 10; Length 367;
 Best Local Similarity 73.3%; Pred. No. 0.0013;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMPYIAGYKTFDGR 15
 DB 102 KMPLYVAGHKTIIDGR 116

RESULT 7
 Q9M4S3 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxId=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257494; AAF72628.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SO SEQUENCE 367 AA; 39832 MW; B5DFP5A61C07A53 CRC64;

Query Match 81.9%; Score 68; DB 10; Length 367;
 Best Local Similarity 73.3%; Pred. No. 0.0013;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMPYIAGYKTFDGR 15
 DB 102 KMPLYVAGHKTIIDGR 116

RESULT 8
 Q9M4S2 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxId=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257495; AAF72629.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SO SEQUENCE 367 AA; 39819 MW; AE7E055A61C07A53 CRC64;

Query Match 81.9%; Score 68; DB 10; Length 367;
 Best Local Similarity 73.3%; Pred. No. 0.0013;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMPYIAGYKTFDGR 15
 DB 102 KMPLYVAGHKTIIDGR 116

RESULT 9
 Q93X51 PRELIMINARY; PRT; 367 AA.
 AC Q93X51;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Putative allergen jun o 1.
 GN JUN O 1.
 OS Juniperus oxycedrus (Prickly juniper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxId=69008;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=POLENT;
 RA "Cloning of Juniperus oxycedrus major allergen."
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ293767; CAC48400.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 SO SEQUENCE 367 AA; 39808 MW; 5D28204DBFD1B9D7 CRC64;

Query Match 81.9%; Score 68; DB 10; Length 367;
 Best Local Similarity 73.3%; Pred. No. 0.0013;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMPYIAGYKTFDGR 15
 DB 102 KMPLYVAGHKTIIDGR 116

RESULT 10
 Q9M4S5 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxId=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257492; AAF72626.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SO SEQUENCE 367 AA; 39894 MW; 5D56FC0E3263B741 CRC64;

Query Match 77.1%; Score 64; DB 10; Length 367;
 Best Local Similarity 66.7%; Pred. No. 0.0064; 1; Indels 0; Gaps 0;
 Matches 10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMPMYIAGYKTPDGR 15
 :||:|||||
 DB 102 EMPYVAGHKTIDGR 116

RESULT 11

Q9A413 PRELIMINARY; PRT; 157 AA.

ID 09A413;
 AC 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE Transcription elongation factor GreA.
 GN CC2848.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gilm M.L., Halt D.H.,
 RA Kojanay J.F., Sait J., Craven M.B., Knouri H., Shetty J., Berry K.,
 RA Uterback T., Tran K., Wolf A., Yamahara J., Ermoleva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AB005950; AAK24812.1; -.
 DR HSBP; P21346; IGRJ.
 DR TIGR; CC2848;
 DR InterPro; IPR001437; GreA_Greb.
 DR Pfam; PF01272; GreA_Greb.1.
 DR Pfam; PF03449; GreA_Greb.N.1.
 DR PRODOM; P0004918; GreA_Greb.1.
 DR PROSITE; PS00829; GREAB.1; UNKNOWN.1.
 KW Elongation factor; Complete proteome.
 SQ SEQUENCE 157 AA; 17232 MW; 880752432BA648D5 CRC64;

Query Match 57.8%; Score 48; DB 16; Length 157;
 Best Local Similarity 61.5%; Pred. No. 1.6;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMPMYIAGYKTPD 13
 :||:|||||
 DB 3 KVMYVAGYOTLD 15

RESULT 12

Q9FM66 PRELIMINARY; PRT; 392 AA.

ID 09FM66;
 AC 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Peccate lyase.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=98290546; PubMed=9628582;

RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 RT Sequence features of the regions of 1,456,315 bp covered by nineteen
 RT physically assigned P1 and TAC clones."
 RL DNA Res. 5:41-54(1998).
 DR EMBL; AB009050; BAB09239.1; -.
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; Pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KW Lyase.

QY 1 KMPMYIAGYKTPDGR 15
 :||:|||||
 DB 129 KQDQVASYTKIDGR 143

Query Match 55.4%; Score 46; DB 10; Length 392;
 Best Local Similarity 53.3%; Pred. No. 9.5;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KMPMYIAGYKTPDGR 15
 :||:|||||
 DB 129 KQDQVASYTKIDGR 143

RESULT 13

Q8RE02 PRELIMINARY; PRT; 122 AA.

ID 08RE02;
 AC 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE Hypothetical protein FN1336.
 GN FN1336.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Valera O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fongstein M., Kyrides N., Overbeek R.,
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 25586."
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AB010638; AAL95532.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 122 AA; 14374 MW; 9EC1D752B888EB20 CRC64;

Query Match 54.2%; Score 45; DB 16; Length 122;
 Best Local Similarity 53.3%; Pred. No. 3.9;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KMPMYIAGYKTPDGR 15
 :||:|||||
 DB 9 KLSKYFGYKPFNGR 23

RESULT 14

Q9VH97 PRELIMINARY; PRT; 3508 AA.

ID 09VH97;
 AC 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE CG9492 protein.
 GN CG9492.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;


```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Burton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake R.G., Champe M., Milos B.D.,
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktiroglu U., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck U., Brokstein P., Broctier P.,
RA Butts K.C., Bussan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Chery J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclob J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Styrkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S.-Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "the genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AE003683; AAF54422.1; -
DR FLYBase; FBgn0037726; CG9492.
DR InterPro; IPR000169; SHProt acsile.
DR PROSITE; PS00639; THIOI. PROTEASE_HIS; UNKNOWN 1.
SQ SEQUENCE 3508 AA; 402200 MW; 235A34368A30D298 CRC64;

Query Match 54.2%; Score 45; DB 5; Length 3508;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMPNYIAGYKTF 12
DB 2954 KLASFIAGYKTF 2965

RESULT 15
O9AVM1 PRELIMINARY; PRT; 357 AA.
AC O9AVM1;
DT 01-UN-2001 (TRENBLREL. 17, Created)
DT 01-UN-2001 (TRENBLREL. 17, last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, last annotation update)
DE mRNA capping enzyme.
GN MCE.
OS Guillardia theta (Cryptomonas phi).
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX NCB1_taxonomy=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20087226; PubMed=10618395;
RA Zauner S., Fraunholz M., Wastl J., Penny S.L., Beaton M.,

```

```

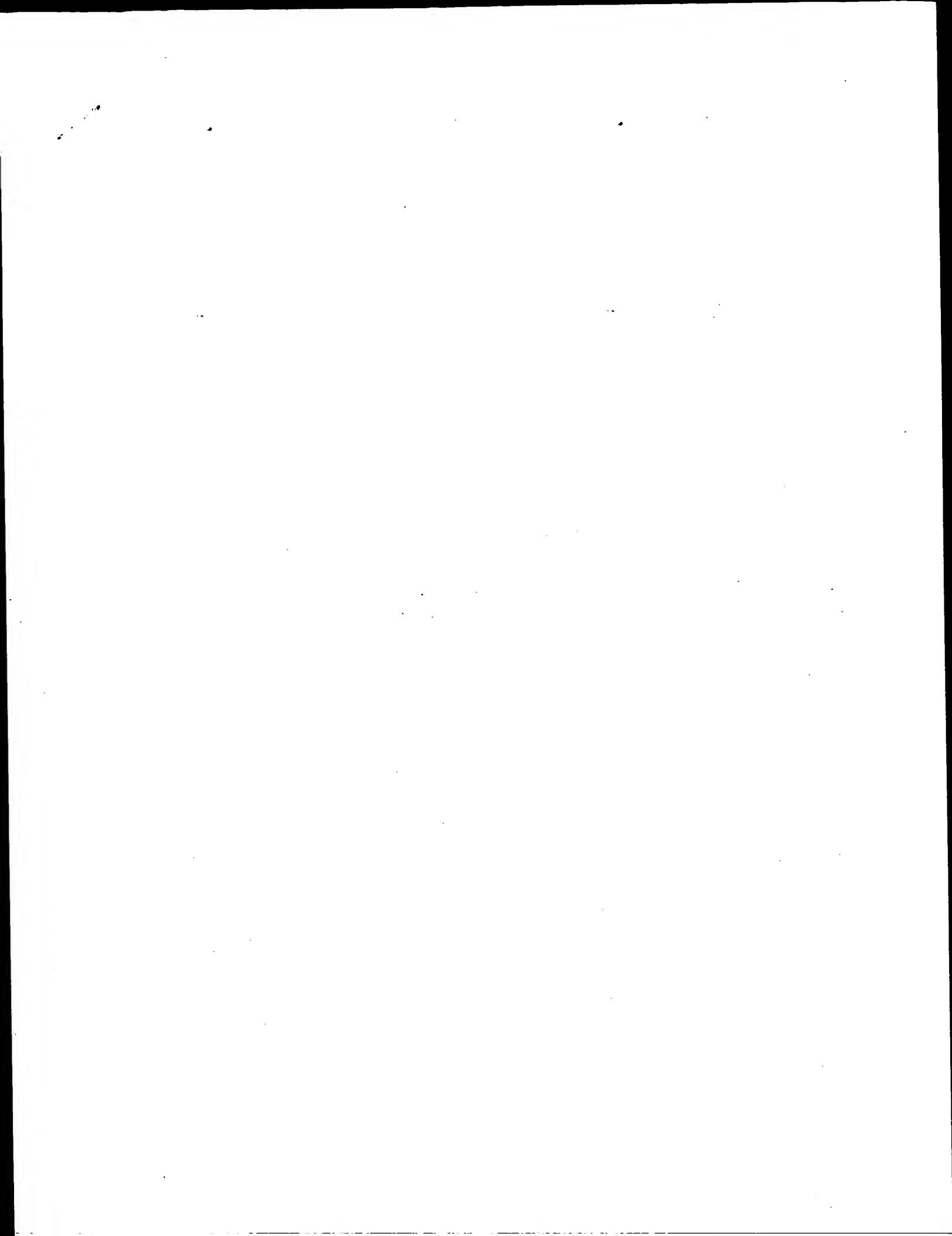
RA Cavalier-Smith T., Maier U., Douglas S.;
RT "Chloroplast protein and centromosomal genes, a tRNA intron, and odd
RT telomeres in an unusually compact eukaryotic genome, the cryptomonad
RT nucleomorph."
RN Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21223349; PubMed=11323671;
RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.,
RA Wu X., Reith M., Cavalier-Smith T., Maier U.;
RT "The highly reduced genome of an enslaved algal nucleus."
RL Nature 410:1091-1096(2001).
DR EMBL; AJ010592; CAC27110.1.
DR InterPro; IPR001339; mRNA_cap_enzyme.
DR Pfam; PF01331; mRNA_cap_enzyme; 1.
SQ SEQUENCE 357 AA; 43120 MW; 6B1BA31F717B15EB CRC64;

Query Match 53.0%; Score 44; DB 10; Length 357;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KMPNYIAGYKTFDG 14
DB 91 KMPNYIAGYKTFDG 104

Search completed: April 20, 2003, 13:13:05
Job time : 14.6711 secs

```



GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds

(without alignments)
237,215 Million cell updates/sec

Title: US-09-142-524d-36

Perfect score: 80
Sequence: 1 PCVPIKRVSNVITIG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	374	2	JC2124
2	80	100.0	374	2	JC2123
3	57	71.2	390	2	H86253
4	56	70.0	394	2	T49115
5	56	70.0	394	2	T49115
6	56	70.0	404	2	T05556
7	55	68.8	404	2	S12209
8	52	65.0	368	2	G86427
9	48	60.0	374	2	T05240
10	48	60.0	374	2	H85148
11	48	60.0	431	2	F86179
12	46	57.5	418	2	T07701
13	44	55.0	463	2	T46165
14	42	52.5	565	2	D84865
15	41	51.2	2647	2	T28161
16	40.5	50.6	528	2	T24160
17	40	50.0	256	2	D84107
18	40	50.0	598	2	T28338
19	40	50.0	2176	2	T39188
20	39	48.8	720	2	S50920
21	39	48.8	961	2	T23395
22	38	47.5	89	2	E53226
23	38	47.5	111	2	A72079
24	38	47.5	111	2	E86544
25	38	47.5	158	2	F65103
26	38	47.5	158	2	B91331
27	38	47.5	158	2	B85976
28	38	47.5	214	2	S44919
29	38	47.5	394	2	G90114

major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar
major allergen Cry j I (clone pCCI-15) #status predicted <MGT>
hypothetical prote
hypothetical prote
pectate lyase like
pectate lyase like
pectate lyase (EC
pectate lyase (EC
Probable pectate 1
pectate lyase (EC
hypothetical prote
hypothetical prote
pectate lyase (EC
pectate lyase-like
hypothetical prote
hypothetical prote
hypothetical prote
ORF MSV077 hypothe
leucic acid ABC
ORF MSV077 hypothe
probable US snRNP
hypothetical prote
retrovirus-related
hypothetical prote
hypothetical prote
PTS system, n-acet
hypothetical prote
hypothetical prote
ZK688.3 protein -
TFIIb related fact

30	38	47.5	587	2	A60367	transforming prote
31	38	47.5	1104	2	T01811	hypothetical prote
32	38	47.5	1328	1	S04273	retrovirus-related
33	37	46.2	71	2	T28314	ORF MSV153 hypothe
34	37	46.2	79	2	B81905	hypothetical prote
35	37	46.2	254	2	E71225	hypothetical prote
36	37	46.2	274	2	T03932	pectate lyase (EC
37	37	46.2	289	2	B32459	pectate lyase (EC
38	37	46.2	289	2	S14235	Na+/K+-exchanging
39	37	46.2	290	2	A34057	Na+/K+-exchanging
40	37	46.2	290	2	A32459	Na+/K+-exchanging
41	37	46.2	290	2	JC5107	Na+/K+-exchanging
42	37	46.2	315	1	C64211	thioredoxin-disulf
43	37	46.2	326	2	S51509	pectate lyase - As
44	37	46.2	327	2	T34400	hypothetical prote
45	37	46.2	421	2	H84466	hypothetical prote

ALIGNMENTS

RESULT 1

major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C/Accession: JC2124

R/Score, T.: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese

A/Reference number: JC2123; PMID:94183234; PMID:8135802

A/Accession: JC2124

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-References: GB:26545; NID:9493633; PIDN:BAA05543.1; PID:9493634

A/Experimental source: Pollen

A/Note: The authors described carbohydrate binding site for residue 279

C/Superfamily: pectate lyase IAT59

C/Keywords: glycoprotein; pollen

F/1-21/Domain: signal sequence #status predicted <SIG>

F/22-374/Product: major allergen Cry j I (clone pCCI-15) #status predicted <MGT>

F/118,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 374;

Best Local Similarity 100.0%; Pred. No. 1.9e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCVPIKRVSNVITIG 15

DB 127 PCVPIKRVSNVITIG 141

RESULT 2

JC2123

major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C/Accession: JC2123; PMID:94183234; PMID:8135802

R/Score, T.: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese

A/Reference number: JC2123; PMID:94183234; PMID:8135802

A/Accession: JC2123

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-References: GB:26544; NID:9493631; PIDN:BAA05542.1; PID:9493632

A/Experimental source: pollen

A/Accession: PC2065

A/Molecule type: protein

A/Residues: 22-53;58-81;219-232;236-258;299-307;346-372 <S02>

A/Note: The authors described carbohydrate binding site for residue 279

C/Superfamily: pectate lyase IAT59

C/Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>
F:2-3/4/Product: major allergen Cry 1 (clone pCCI-2-2) #status predicted <MAT>
F:158,191,293,354/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 100.0%; Score 80; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCVFIRKRVSNVTHG 15
DB 127 PCVFIRKRVSNVTHG 141

RESULT 3

H86253
hypothetical,protein (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H86253
R:Theologian, A.; Becker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A6141; MUID:21016719; PMID:11130712
A:Accession: H86253
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-390 <STO>
A:Cross-references: GB:AE005172; NID:G3157942; PIDN:AA17625.1; GSPDB:GN00141
A:Genetics:
A:Map position: 1
C:Superfamily: pectate lyase LAT59

Query Match 71.2%; Score 57; DB 2; Length 390;
Best Local Similarity 66.7%; Pred. No. 0.026;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 PCVFIRKRVSNVTHG 15
DB 140 PCITIQVTVNTHG 154

RESULT 4

T49115
pectate lyase like protein - Arabidopsis thaliana
N:Alternate names: protein AT4g22080
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000
C:Accession: T49115
R:Bevan, M.; Medler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25016
A:Accession: T49115
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <BEV>
A:Cross-references: EMBL:AL022140; GSPDB:GN00062; ATSP:AT4g22080
A:Experimental source: cultivar Columbia; BAC clone FIN20
C:Genetics:
A:Gene: ATSP:AT4g22080
A:Map position: 4
A:Introns: 240/3; 275/2; 321/3
C:Superfamily: pectate lyase LAT59

Query Match 70.0%; Score 56; DB 2; Length 394;
Best Local Similarity 60.0%; Pred. No. 0.04;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 PCVFIRKRVSNVTHG 15
DB 150 PCITIQVTVNTHG 164

RESULT 5

T49116
pectate lyase like protein - Arabidopsis thaliana
N:Alternate names: protein AT4g22090
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C:Accession: T49116
R:Bevan, M.; Medler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25016
A:Accession: T49116
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <BEV>
A:Cross-references: EMBL:AL022140; GSPDB:GN00062; ATSP:AT4g22090
A:Experimental source: cultivar Columbia; BAC clone FIN20
C:Genetics:
A:Gene: ATSP:AT4g22090
A:Map position: 4
A:Introns: 240/3; 275/2; 321/3
C:Superfamily: pectate lyase LAT59

Query Match 70.0%; Score 56; DB 2; Length 394;
Best Local Similarity 60.0%; Pred. No. 0.04;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 PCVFIRKRVSNVTHG 15
DB 150 PCITIQVTVNTHG 164

RESULT 6

T05556
pectate lyase (EC 4.2.2.2) F22K18.20 - Arabidopsis thaliana
N:Alternate names: protein F22K18.20
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05556
R:Bevan, M.; Medler, H.; Wambutt, R.; Hohnsels, J.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15419
A:Accession: T05556
A:Molecule type: DNA
A:Residues: 1-404 <BEV>
A:Cross-references: EMBL:AL035356
A:Experimental source: cultivar Columbia; BAC clone F22K18
C:Genetics:
A:Map position: 4
A:Introns: 37/2; 287/2
A:Note: F22K18.20
C:Superfamily: pectate lyase LAT59
C:Keywords: carbon-oxygen lyase

Query Match 70.0%; Score 56; DB 2; Length 404;
Best Local Similarity 60.0%; Pred. No. 0.041;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 PCVFIRKRVSNVTHG 15
DB 162 PCITIQVTVNTHG 176

RESULT 7

S12209
pectate lyase (EC 4.2.2.2) - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 23-Jul-1999

C:Accession: S12209
R:Budellier, K.A.; Smith, A.G.; Gasser, C.S.
Mol. Gen. Genet. 224, 183-192, 1990
A>Title: Regulation of a stylar transmitting tissue-specific gene in wild-type and trans
A:Reference number: S12209; MUID:91117185; PMID:2277637
A:Accession: S12209
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-404 <BD>
A:Cross-references: GB:X55193; NID:919161; PIDN:CAA8979.1; PID:919162
C:Superfamily: pectate lyase LAT59
C:Keywords: carbon-oxygen lyase

Query Match
Best Local Similarity 60.0%; Score 55; DB 2; Length 404;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 158 PCIRIRKVNVIING 172

RESULT 8
G86427
probable pectate lyase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G86427
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, U.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matli, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86427
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-368 <STO>
A:Cross-references: GB:AE005172; NID:91055821; PIDN:AAG28291.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: pectate lyase LAT59

Query Match
Best Local Similarity 65.0%; Score 52; DB 2; Length 368;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 139 PCIRIRKVNVIING 153

RESULT 9
T05240
pectate lyase (EC 4.2.2.2) A11 - Arabidopsis thaliana (fragment)
N/Alternate names: protein F18A5.100
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05240
R:Bevan, M.; Weber, N.; Grueninger, D.; Schmidheini, T.; Bancroft, I.; Mewes, H.W.; Maye
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15405
A:Accession: T05240
A:Molecule type: DNA
A:Residues: 1-374 <BEV>
A:Cross-references: EMBL:AL035528
C:Genetics:
A:Experimental source: cultivar Columbia, BAC clone F18A5

A:Map position: 4
A>Note: Intron positions not resolved
A>Note: F18A5.100
C:Superfamily: pectate lyase LAT59
C:Keywords: carbon-oxygen lyase

Query Match
Best Local Similarity 60.0%; Score 48; DB 2; Length 374;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 131 CITIOTPTNVIING 144

RESULT 10
H85148
probable pectate lyase A11 (partial) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: H85148
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spr
Nature 402, 769-777, 1999
A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85148
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <STO>
A:Cross-references: GB:NC_001268; NID:97268075; PIDN:CAB78413.1; GSPDB:GN00140
C:Genetics:
A:Gene: ATG413710
A:Map position: 4
C:Superfamily: pectate lyase LAT59

Query Match
Best Local Similarity 60.0%; Score 48; DB 2; Length 374;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 131 CITIOTPTNVIING 144

RESULT 11
F86179
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F86179
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, U.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matli, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86179
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-431 <STO>
A:Cross-references: GB:AE005172; NID:92494113; PIDN:AAB80622.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: pectate lyase LAT59

Query Match
Best Local Similarity 60.0%; Score 48; DB 2; Length 431;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 CVFIKRVSNVITIG 15
 DB 186 CITIQVYVNIHIG 199

RESULT 12

T07701
 pectate lyase (EC 4.2.2.2) F17N18.100 - Arabidopsis thaliana
 N:Alternate names: protein F17N18.100
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Jul-1999
 C:Accession: T07701
 R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueb
 submitted to the Protein Sequence Database, May 1999
 A:Reference number: 216098
 A:Accession: T07701
 A:Molecule type: DNA
 A:Residues: 1-418 <BEV>
 A:Cross-references: EMBL:AL049751
 A:Experimental source: cultivar Columbia; BAC clone F17N18
 C:Genetics:
 A:Map position: 4
 A:introns: 26/2; 49/2; 264/3; 299/2; 345/3
 A:Note: F17N18.100
 C:Superfamily: pectate lyase LATS9
 C:Keywords: carbon-oxygen lyase

Query Match 57.5%; Score 46; DB 2; Length 418;
 Best Local Similarity 50.0%; Pred. No. 2.6;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 CVFIKRVSNVITIG 15
 DB 175 CITVQVYVNIHIG 188

RESULT 13

T46165
 pectate lyase-like protein - Arabidopsis thaliana
 N:Alternate names: protein T4D2.120
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
 R:Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Welcheslgaertner, M.;
 submitted to the Protein Sequence Database, December 1999
 A:Reference number: 223025
 A:Accession: T46165
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-463 <NTA>
 A:Cross-references: EMBL:AL132958
 A:Experimental source: cultivar Columbia; BAC clone T4D2
 C:Genetics:
 A:Map position: 3
 A:introns: 27/2; 112/2; 182/1; 265/3; 329/3; 399/2
 A:Note: T4D2.120
 C:Superfamily: pectate lyase LATS9

Query Match 55.0%; Score 44; DB 2; Length 463;
 Best Local Similarity 53.8%; Pred. No. 6.6;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 CVFIKRVSNVITIG 14
 DB 159 CITIQVYVNIHIG 171

RESULT 14

D84886
 hypothetical protein At2g45100 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: D84886
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Yankken, S.E.; Umayam, L.; Tallon, L.;
 eus, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84886
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-565 <STO>
 A:Cross-references: GB:AE002093; NID:94895242; PIDN:ADJ32827.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g45100
 A:Map position: 2

Query Match 52.5%; Score 42; DB 2; Length 565;
 Best Local Similarity 46.7%; Pred. No. 18;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 PCVFIKRVSNVITIG 15
 DB 176 PSIFPFRSMNLIG 190

RESULT 15

T28161
 hypothetical protein FCR3-varT11-1 - malaria parasite (Plasmodium falciparum) (fragments)
 C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T28161
 R:Hernandez-Rivas, R.; Mattei, D.; Sterkers, Y.; Peterson, D.S.; Wellens, T.E.; Scherf,
 Mol. Cell. Biol. 17, 604-611, 1997
 A:Title: Expressed var genes are found in Plasmodium falciparum subtelomeric regions.
 A:Reference number: Z20483; MUID:97154495; PMID:9001213
 A:Accession: T28161
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2647 <HBR>
 A:Cross-references: EMBL:U67959; NID:91794255; PID:G1809295; PIDN:AAC47438.1
 A:Experimental source: strain FQ27/PNG
 C:Genetics:
 A:Note: FCR3-varT11-1
 A:introns: 2158/3

Query Match 51.2%; Score 41; DB 2; Length 2647;
 Best Local Similarity 42.9%; Pred. No. 1.2e+02;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 CVFIKRVSNVITIG 15
 DB 1070 CKRVKIKKMLIIG 1083

Search completed: April 20, 2003, 13:15:46
 Job time : 8.07895 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05; Search time 18.1974 Seconds
(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524d-121

Perfect score: 79

Sequence: 1 ASKNFHLQKNTIGTG 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	15	17	AAK97908
2	79	100.0	15	19	AAW57764
3	79	100.0	19	22	AAAB4115
4	79	100.0	42	22	AAAB4105
5	79	100.0	80	18	AAW27369
6	79	100.0	105	18	AAW27370
7	79	100.0	134	18	AAW27371
8	79	100.0	460	16	AAK69791
9	79	100.0	514	15	AAK53690
10	79	100.0	514	16	AAK74333

11	79	100.0	514	16	AAK69792	Japonicum allergen
12	79	100.0	514	17	AAK93589	Japan cedar pollen
13	79	100.0	514	17	AAK81586	Cedar pollen aller
14	79	100.0	514	20	AAZ25666	Japanese cedar all
15	79	100.0	514	20	AAZ25667	Japanese cedar all
16	73	92.4	17	17	AAK81582	Japonicum allergen
17	68	86.1	17	19	AAW80347	Sugi allergen aller
18	59	74.7	47	19	AAW80351	Sugi allergen prot
19	59	74.7	47	19	AAW80351	Sugi allergen prot
20	59	74.7	54	21	AAK23905	Artificial sequenc
21	59	74.7	54	21	AAK23905	Artificial sequenc
22	59	74.7	74	21	AAK23905	Artificial sequenc
23	59	74.7	94	21	AAK23905	Artificial sequenc
24	59	74.7	95	21	AAK23901	Artificial sequenc
25	59	74.7	95	21	AAK23901	Artificial sequenc
26	59	74.7	95	22	AAK69098	Cedar pollen aller
27	59	74.7	96	22	AAK69104	Cedar pollen aller
28	59	74.7	96	22	AAK69105	Cedar pollen aller
29	59	74.7	97	22	AAK69105	Cedar pollen aller
30	59	74.7	99	21	AAK23902	Artificial sequenc
31	59	74.7	99	21	AAK23907	Artificial sequenc
32	59	74.7	99	22	AAK69099	Cedar pollen aller
33	59	74.7	214	22	AAK69120	Cedar pollen aller
34	58	73.4	12	18	AAW12549	Japanese cedar pol
35	58	73.4	12	22	AAK69108	Cryptomeria japon
36	58	73.4	47	19	AAW80352	Sugi allergen prot
37	58	73.4	47	19	AAW80353	Sugi allergen prot
38	58	73.4	47	19	AAW80356	Sugi allergen prot
39	58	73.4	47	19	AAW80357	Sugi allergen prot
40	58	73.4	61	21	AAK23883	Artificial sequenc
41	58	73.4	61	21	AAK23883	Artificial sequenc
42	58	73.4	61	21	AAK23894	Artificial sequenc
43	58	73.4	71	21	AAK23878	Artificial sequenc
44	58	73.4	79	21	AAK23880	Artificial sequenc
45	58	73.4	79	21	AAK23884	Artificial sequenc

ALIGNMENTS

RESULT 1
AAK97908
ID AAK97908 standard; peptide; 15 AA.
XX
AC AAK97908;
XX
XX 16-AUG-1996 (first entry)
XX
XX Japan cedar pollen mature allergen Cry j II amino acids 186-200.
XX
XX Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
XX Sugi pollinosis; diagnosis; treatment.
XX
XX Cryptomeria japonica.
XX
XX JP08047392-A.
XX
XX 20-FEB-1996.
XX
XX
XX 07-NOV-1994; 94UP-0297840.
XX
XX 26-MAY-1994; 94UP-0134868.
XX
XX 05-NOV-1993; 93UP-0276773.
XX
XX (MEIP) MEIJI MILK PROD CO LTD.
XX
XX WPI; 1996-166249/17.
XX
XX Japan cedar pollen allergen Cry j II epitope - comprises at least
XX part of specified 460 amino acid protein
XX
XX Claim 8; Fig 4; 17pp; Japanese.
XX

CC AAR97871-R97960 are overlapping peptides used for the epitope mapping
 CC of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic
 CC peptides of it are useful in the diagnosis, prevention and treatment
 CC of Sugi pollinosis, the allergic reaction to Japan cedar pollen.
 CC Significant regions of the allergen were identified using the
 CC overlapping peptides of the full epitope derived from a Cry j II
 CC antigen-specific T cell line. Amino acids 66-80 (AAR97884) and 186-200
 CC (R978908) of the full mature 460 amino acid allergen are the most
 CC allergenic of the 90 peptides tested.

XX Sequence 15 AA;

Query Match 100.0%; Score 79; DB 17; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.3e-07; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNFHLQKNTIGTG 15
 DB 1 ASKNFHLQKNTIGTG 15

RESULT 2

AAM57764

ID AAM57764 standard; peptide; 15 AA.

AC AAM57764;

DT 17-SEP-1998 (first entry)

DE Residues 186-200 of Cry j 2.

KW Cry j 2; Japanese cedar pollen antigen; allergy; immunotherapy;

KM HLA class II molecule.

OS Cryptomeria japonica.

PN WO9820902-A1.

PD 22-MAY-1998.

PF 12-NOV-1997; 97WO-JP04129.

PR 13-NOV-1996; 96JP-0302053.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Kino K, Kume A, Sone T;

DR WPI; 1998-297617/26.

PT Peptides derived from Japanese cedar pollen antigens are
 PT immunotherapeutic agents - useful for allergy treatment and typing
 PT HLA class II molecules in allergy sufferers

PS Claim 12; Page 31; 50pp; Japanese.

CC This sequence represents residues 186-200 of the Cry j 2 protein, and
 CC is a peptide of the invention. The peptides are derived from Japanese
 CC cedar pollen antigens, and are used as immunotherapeutic agents in the
 CC treatment of allergy. The peptides can be used for identification and
 CC typing of the particular HLA class II molecules in an allergy sufferer,
 CC and also for peptide immunotherapy of an allergy. Using these peptides
 CC the immunotherapy can be targeted more specifically to the requirements
 CC of the individual patient, allowing more effective treatment of an
 CC allergy, including those patients for whom treatment with a conventional
 CC immunotherapeutic agent is ineffective.

XX Sequence 15 AA;

Query Match 100.0%; Score 79; DB 19; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.3e-07; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNFHLQKNTIGTG 15
 DB 1 ASKNFHLQKNTIGTG 15

RESULT 3

AAB84115

ID AAB84115 standard; peptide; 19 AA.

AC AAB84115;

DT 06-AUG-2001 (first entry)

DE Amino acid sequence of a pollen antigen.

KW Immunomodulatory peptide; immune response; autoimmune disease; allergy;

KM asthma; host-venous-graft rejection; T cell; anergy; apoptosis.

OS Cryptomeria japonica.

PN WO200136448-A2.

PD 25-MAY-2001.

PF 27-OCT-2000; 2000WO-US41646.

PR 27-OCT-1999; 99US-0161734.

PA (CELS-) CEL-SCI CORP.

PI Zimmerman DH;

DR WPI; 2001-374498/39.

PT Novel immunomodulatory peptide construct useful for modulating an
 PT inappropriate immune response in an individual at risk for autoimmune
 PT disease, allergic reactions, asthma or host-graft or graft-host disease

PS Claim 2; Page 27; 55pp; English.

CC The specification describes an immunomodulatory peptide construct.
 CC The immunomodulatory peptide comprises a first peptide associated with
 CC autoimmune disease, allergy or asthma, or host-venous-graft rejection
 CC and which will bind to an antigen receptor on a set or subset of T cells,
 CC linked to a second immune modulating peptide which will cause a directed
 CC immune response by the set or subset of T cells to which the first
 CC peptide is attached. Alternatively, the second peptide will bind to a
 CC T cell receptor site on the surface of the T cell which will cause the
 CC set or subset of T cells to initiate, but not complete, an immune
 CC response to cause the set or subset of T cells to undergo anergy and
 CC apoptosis. The immunomodulatory peptides are useful for eliminating a set
 CC or subset of T cells involved in autoimmune response. They are useful for
 CC the treatment of autoimmune disease, allergic reactions, asthma or
 CC host-graft or graft-host rejections. The immunomodulatory peptides are
 CC also useful for interrupting an autoimmune disease associated pathway
 CC necessary to complete T cell activation. The present peptide is used to
 CC construct immunomodulatory peptides of the invention, and is
 CC representative of the first peptide.

XX Sequence 19 AA;

Query Match 100.0%; Score 79; DB 22; Length 19;

Best Local Similarity 100.0%; Pred. No. 1.7e-07; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNFHLQKNTIGTG 15
 DB 5 ASKNFHLQKNTIGTG 19

RESULT 4

AAB84105

ID AAB84105 standard; peptide; 42 AA.
 AC AAB84105;
 XX
 DT 06-AUG-2001 (first entry)
 XX
 DE Immunomodulatory peptide used to treat allergic responses.
 XX
 DE Immunomodulatory peptide; immune response; autoimmune disease; allergy;
 KW asthma; host-versus-graft rejection; T cell; anergy; apoptosis.
 XX
 OS Synthetic.
 OS Cryptomeria japonica.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..17
 FT Peptide /note= "LPA-3 peptide"
 FT Peptide 18..23
 FT Peptide /note= "linker"
 FT Peptide 24..42
 FT Peptide /note= "pollen antigen"
 XX
 PN WO200136448-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 27-OCT-2000; 2000MO-US41646.
 XX
 PR 27-OCT-1999; 99US-0161734.
 XX
 PA (CELS-) CEL-SCT CORP.
 XX
 PI Zimmerman DH;
 XX
 FI WPI; 2001-374498/39.
 DR
 XX
 PT Novel immunomodulatory peptide construct useful for modulating an
 PT inappropriate immune response in an individual at risk for autoimmune
 PT disease, allergic reactions, asthma or host-graft or graft-host disease
 PT
 XX
 PS Claim 4; Page 27; 55pp; English.
 XX
 CC The specification describes an immunomodulatory peptide construct.
 CC The immunomodulatory peptide comprises a first peptide associated with
 CC autoimmune disease, allergy or asthma, or host-versus-graft rejection
 CC and which will bind to an antigen receptor on a set or subset of T cells,
 CC linked to a second immune modulating peptide which will cause a directed
 CC immune response by the set or subset of T cells to which the first
 CC peptide is attached. Alternatively, the second peptide will bind to a
 CC T cell receptor site on the surface of the T cell which will bind to a
 CC set or subset of T cells to initiate, but not complete, an immune
 CC response to cause the set or subset of T cells to undergo anergy and
 CC apoptosis. The immunomodulatory peptides are useful for eliminating a set
 CC or subset of T cells involved in autoimmune response. They are useful for
 CC the treatment of autoimmune disease, allergic reactions, asthma or
 CC host-graft or graft-host rejections. The immunomodulatory peptides are
 CC also useful for interrupting an autoimmune disease associated pathway
 CC necessary to complete T cell activation. The present sequence represents
 CC an immunomodulatory peptide used to treat allergic responses.
 XX
 XX
 SQ Sequence 42 AA;
 XX
 Query Match 100.0%; Score 79; DB 22; Length 42;
 Best Local Similarity 100.0%; Pred. No. 4e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASKNFHLQKNTIGTG 15
 DB 28 ASKNFHLQKNTIGTG 42

RESULT 5
 ID AAM27369 standard; peptide; 80 AA.
 AC AAM27369;
 XX
 DT 24-MAR-1998 (first entry)
 XX
 DE Multi-epitope peptide used as immunotherapeutic agent #1.
 XX
 DE Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 XX
 OS Synthetic.
 OS WO9732600-A1.
 XX
 PN 12-SEP-1997.
 XX
 PF 10-MAR-1997; 97WO-JP00740.
 XX
 PR 10-MAR-1996; 96JP-0080702.
 XX
 PA (MEIP) MEIJI MILK PROD CO LTD.
 XX
 PI Dairiki K, Iwama A, Kuno K, Kume A, Sone T;
 XX
 FI WPI; 1997-470495/43.
 DR
 XX
 PT Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens
 XX
 PS Claim 6; Page 31; 58pp; Japanese.
 XX
 CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX
 SQ Sequence 80 AA;
 XX
 Query Match 100.0%; Score 79; DB 18; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASKNFHLQKNTIGTG 15
 DB 35 ASKNFHLQKNTIGTG 49

RESULT 6
 ID AAM27370 standard; peptide; 105 AA.
 AC AAM27370;
 XX
 DT 24-MAR-1998 (first entry)
 XX
 DE Multi-epitope peptide used as immunotherapeutic agent #2.
 XX
 DE Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 XX
 OS Synthetic.
 OS WO9732600-A1.
 PN

XX 12-SEP-1997.
 PD 10-MAR-1997; 97WO-JP00740.
 PF 10-MAR-1996; 96UP-0080702.
 XX (MEIP) MEIJI MILK PROD CO LTD.
 PA Dairiki K, Iwama A, Kino K, Kume A, Sone T;
 PI WPI; 1997-470495/43.
 DR Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens
 PS Claim 6; Page 31; 58pp; Japanese.
 XX The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX Sequence 105 AA;

Query Match 100.0%; Score 79; DB 18; Length 105;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ASKNFHLQKNTIGTG 15
 DB 35 ASKNFHLQKNTIGTG 49

RESULT 7
 ID AAW27371 standard; peptide; 134 AA.
 XX AAW27371;
 AC AAW27371;
 DT 24-MAR-1998. (first entry)
 DE Multi-epitope peptide used as immunotherapeutic agent #3.
 XX Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 OS Synthetic.
 XX WO9732600-A1.
 PN 12-SEP-1997.
 PD 10-MAR-1997; 97WO-JP00740.
 PR 10-MAR-1996; 96UP-0080702.
 XX (MEIP) MEIJI MILK PROD CO LTD.
 PA Dairiki K, Iwama A, Kino K, Kume A, Sone T;
 PI WPI; 1997-470495/43.
 DR Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens

XX Claim 6; Page 32; 58pp; Japanese.
 PS The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX Sequence 134 AA;

Query Match 100.0%; Score 79; DB 18; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ASKNFHLQKNTIGTG 15
 DB 35 ASKNFHLQKNTIGTG 49

RESULT 8
 ID AAR69791 standard; Protein; 460 AA.
 XX AAR69791;
 AC AAR69791;
 DT 27-SEP-1995 (first entry)
 DE Japonicum allergen residues 55-514.
 XX Japonicum allergen; residues 55-514; induced histamine release;
 KW antiallergic peptide; IgE cross-linking inhibition.
 XX Japonicum sp.
 OS WO9502412-A.
 PN 26-JAN-1995.
 PD 15-JUL-1994; 94WO-JP01164.
 PF 16-JUL-1993; 93JP-0177008.
 PR 01-SEP-1993; 93JP-0217725.
 PR 07-APR-1994; 94JP-0069336.
 XX (MEIP) MEIJI MILK PROD CO LTD.
 PA Kino K, Kohno Y, Komiyama N, Sone T;
 PI WPI; 1995-067159/09.
 DR N-PSDB; AAQ84044.
 PT Peptide antiallergic agent - inhibits cross-linking of allergen
 PT with IgE antibody
 PS Disclosure; Pages 26-27; 46pp; Japanese.
 XX AAQ84044 encodes AAR69791 Japonicum allergen residues 55-514, from
 CC which the antiallergic peptides AAR69845-R69809 were derived.
 CC The peptides ability to inhibit the cross-linking of an allergen,
 CC to an IgE antibody can be used in the prevention and treatment of
 CC allergic diseases.
 XX Sequence 460 AA;

Query Match 100.0%; Score 79; DB 16; Length 460;
 Best Local Similarity 100.0%; Pred. No. 5.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNFHLOKNTTGTG 15
DB 186 ASKNFHLOKNTTGTG 200

RESULT 9

AAR53690
ID AAR53690 standard; Protein; 514 AA.
XX
AC AAR53690;

DT 01-FEB-1995 (first entry)

DE Japanese cedar pollen allergen Cry j II.

XX Cedar pollinosis; diagnostic.

OS Cryptomeria japonica.

PN WO9411512-A.

PD 26-MAY-1994.

PF 12-NOV-1993; 93WO-US11000.

PR 12-NOV-1992; 92US-0975179.

PA (IMMU-)* IMMUNOLOGIC PHARM CORP.

PI Brauer A, Kuo M, Pollock J, Yeung S;

DR WPI; 1994-183513/22.

DR N-PSDB; AAO66048.

PT Allergenic Cry j II protein and fragments from Japanese cedar
pollen - used to diagnose, treat and prevent Japanese cedar
pollinosis

PS Claim 2; Fig 4; 89pp; English.

CC The sequence is of a Japanese cedar pollen allergen Cry j
II. The protein and its fragments can be used for diagnosis and
treatment of Japanese cedar pollinosis and to identify similar
sequences in other plants.
CC See also AAR53692-6.

CC Sequence 514 AA;

Query Match Best Local Similarity 100.0%; Score 79; DB 15; Length 514;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNFHLOKNTTGTG 15
DB 240 ASKNFHLOKNTTGTG 254

RESULT 10

AAR74333
ID AAR74333 standard; Protein; 514 AA.

AC AAR74333;

DT 01-NOV-1995 (first entry)

DE Japanese cedar pollen allergen.

XX Japanese cedar; pollen; allergen; allergy; therapy; diagnostic;

OS Cryptomeria japonica.

PN EP655500-A.

PD 31-MAY-1995.

PF 03-NOV-1994; 94EP-0308117.

PR 27-DEC-1993; 93JP-0346814.

PR 05-NOV-1993; 93JP-0292151.

PR 20-DEC-1993; 93JP-0344596.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

PI Kurimoto M, Namba M, Torigoe K;

DR WPI; 1995-195588/26.

DR N-PSDB; AAO90156.

PT New Japanese cedar pollen allergen polypeptide - and DNA coding for it,
useful for treatment and diagnosis of cedar pollen allergy

XX Claim 5; Page 26-28; 41pp; English.

CC The gene encoding an allergen of Japanese cedar pollen was isolated
by PCR amplification using primers based on portions of the allergen
protein. The gene was used for recombinant allergen production in
E. coli (vector plasmid pKK-223-3).

CC Sequence 514 AA;

Query Match Best Local Similarity 100.0%; Score 79; DB 16; Length 514;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNFHLOKNTTGTG 15
DB 240 ASKNFHLOKNTTGTG 254

RESULT 11

AAR69792
ID AAR69792 standard; Protein; 514 AA.

AC AAR69792;

DT 27-SEP-1995 (first entry)

DE Japonicum allergen.

KM Japonicum allergen; induced histamine release; antiallergic peptide;
IgE cross-linking inhibition.

XX Japonicum sp.

OS WO9502412-A.

PN 26-JAN-1995.

PF 15-JUL-1994; 94WO-JP01164.

PR 16-JUL-1993; 93JP-0177008.

PR 01-SEP-1993; 93JP-0217725.

PR 07-APR-1994; 94JP-0069336.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Kino K, Kohno Y, Komiyama N, Sone T;

DR WPI; 1995-067159/09.
DR N-PSDB; AAO84045, AAO84046.
PT Peptide antiallergic agent - inhibits cross-linking of allergen
with IgE antibody

PS Example 3; Pages 27-28; 46pp; Japanese.

XX AA084045 encodes AAR69792 Japonicum allergen, from which the
CC anti-allergic peptides AAR6945-R69809 were derived. The peptides ability
CC to inhibit the cross-linking of an allergen, to an IgE antibody can be
CC used in the prevention and treatment of allergic diseases.

XX Sequence 514 AA;

Query Match 100.0%; Score 79; DB 16; Length 514;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNFHLQKNTIGTG 15
DB 240 ASKNFHLQKNTIGTG 254

RESULT 12
AAR93599
ID AAR93599 standard; Protein; 514 AA.

XX AC AAR93599;

XX DT 16-AUG-1996 (first entry)

XX DE Japan cedar pollen Cry j II allergen.

XX Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
XX Sugi pollinosis; diagnosis; treatment.

XX OS Cryptomeria japonica.

XX FH Key Location/Qualifiers

XX FT Peptide 1..54

XX FT /*label= sig_peptide

XX FT 55..514

XX FT /*label= mat_protein

XX PN JP08047392-A.

XX PD 20-FEB-1996.

XX PP 07-NOV-1994; 94JP-0297840.

XX PR 26-MAY-1994; 94JP-0134868.

XX PR 05-NOV-1993; 93JP-0276773.

XX XX (WEIP) MEIUI MILK PROD CO LTD.

XX PA

XX DR WPI; 1996-166249/17.

XX DR N-PSDB; AAT18102.

XX FT Japan cedar pollen allergen Cry j II epitope - comprises at least

XX PT part of specified 460 amino acid protein

XX PS Claim 1; Page 10-11; 17pp; Japanese.

XX CC AAR93599 is a Japan cedar pollen Cry j II allergen which is useful

XX CC in the diagnosis, prevention and treatment of Sugi pollinosis.

XX CC the allergic reaction to Japan cedar pollen. Significant regions of

XX CC the allergen were identified using overlapping peptides of the full

XX CC epitope derived from a Cry j II antigen-specific T cell line

XX CC (see AAR97871-R97960). Amino acids 66-80 (AAR97884) and 186-200

XX CC (R978908) of the full mature 460 amino acid allergen are the most

XX CC allergenic of the 90 peptides tested.

XX SQ Sequence 514 AA;

Query Match 100.0%; Score 79; DB 17; Length 514;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNFHLQKNTIGTG 15
DB 240 ASKNFHLQKNTIGTG 254

RESULT 13

XX AAR1586
ID AAR1586 standard; Protein; 514 AA.

XX AC AAR1586;

XX DT 24-MAY-1996 (first entry)

XX DE Cedar pollen allergen A.

XX Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
XX antibody; pollinosis; therapy; immunotherapy.

XX OS Cryptomeria japonica.

XX PN EP700929-A2.

XX PD 13-MAR-1996.

XX PF 08-SEP-1995; 95EP-0306295.

XX PR 14-JUL-1995; 95JP-0200221.

XX PR 10-SEP-1994; 94JP-0242137.

XX PR 14-JUL-1995; 95JP-0200204.

XX PA (HAYB) HAYASHIBARA SEIUTSU KAGAKU.

XX PI Hino K, Saito S, Taniguchi Y;

XX DR WPI; 1996-140976/15.

XX PT New peptide(s) derived from cedar pollen allergens - activate

XX PT allergen-specific T-cells, but not allergen-specific IgE antibodies,

XX PT used for treating cedar pollinosis

XX PS Claim 5; Page 29-30; 36pp; English.

XX CC Synthetic peptides based on portions of cedar pollen allergens A

XX CC (AAR81586) and B (AAR81587) were tested for their ability to activate

XX CC cedar allergen-specific T-cells, but not allergen-specific IgE

XX CC antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell

XX CC epitopes. These peptides, plus subsequences (AAR81573-79) essential

XX CC for T-cell recognition, and homologous peptides (AAR81588-96) can

XX CC be used as immunotherapeutic agents to treat or prevent cedar

XX CC pollinosis, avoiding side-effects such as anaphylaxis.

XX SQ Sequence 514 AA;

Query Match 100.0%; Score 79; DB 17; Length 514;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNFHLQKNTIGTG 15
DB 240 ASKNFHLQKNTIGTG 254

RESULT 14

XX AAR25666
ID AAR25666 standard; Protein; 514 AA.

XX AC AAR25666;

XX DT 30-SEP-1999 (first entry)

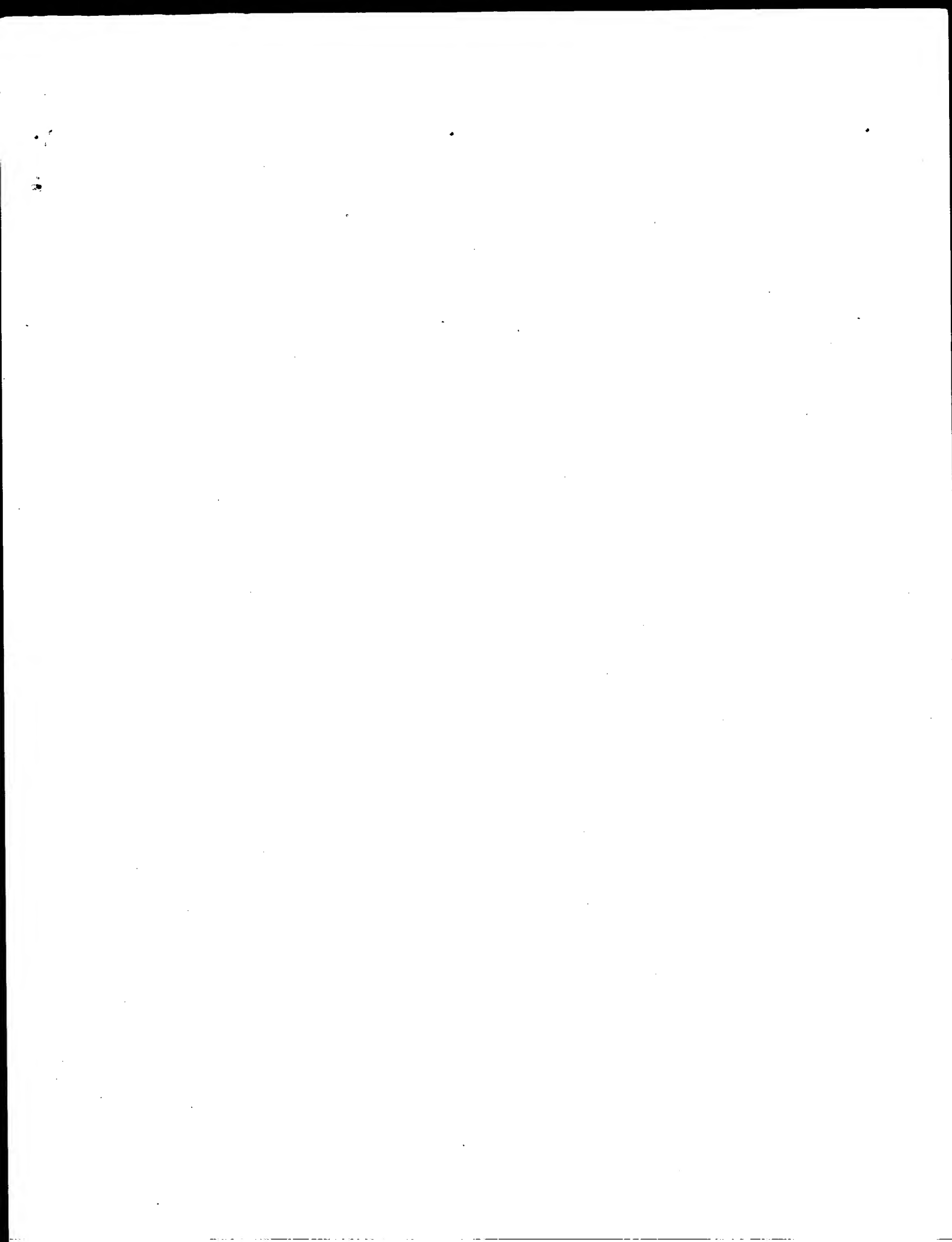
XX DE Japanese cedar allergen 1076242 Cry j II precursor protein fragment.

XX XX Major histocompatibility complex; class II; desensitising; human;

KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; stinging;
 KW chitoniidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 OS Cedrus sp.
 XX WO9934826-A1.
 XX 15-JUL-1999.
 XX 11-JAN-1999; 99WO-GB00080.
 XX 21-SEP-1998; 98GB-0020474.
 XX 09-JAN-1998; 98GB-0000445.
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PI Kay AB, Larche M;
 DR WPI; 1999-458255/38.
 PT Desensitizing patients to polypeptide allergens
 PS Example 6; Page 74; 117pp; English.
 XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chitoniidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrilio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076242 Cry j II
 CC precursor.
 XX
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 79; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 6,7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASKNFHLQKNTIGTG 15
 DB 240 ASKNFHLQKNTIGTG 254
 RESULT 15
 AAY25667
 ID AAY25667 standard; protein; 514 AA.
 AC AAY25667;
 XX 30-SEP-1999 (first entry)
 DT Japanese cedar allergen 1076241 Cry j II protein fragment.
 DE
 XX Major histocompatibility complex; class II; desensitizing; human;
 KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; stinging;
 KW chitoniidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX

OS Cedrus sp.
 XX WO9934826-A1.
 XX 15-JUL-1999.
 XX 11-JAN-1999; 99WO-GB00080.
 XX 21-SEP-1998; 98GB-0020474.
 XX 09-JAN-1998; 98GB-0000445.
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PI Kay AB, Larche M;
 DR WPI; 1999-458255/38.
 PT Desensitizing patients to polypeptide allergens
 PS Example 6; Page 74; 117pp; English.
 XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chitoniidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrilio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076241 Cry j II.
 XX
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 79; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 6,7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASKNFHLQKNTIGTG 15
 DB 240 ASKNFHLQKNTIGTG 254

Search completed: April 20, 2003, 13:06:11
 Job time : 18.1974 secs



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds

(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524d-131

Perfect score: 76

Sequence: 1 SRAEVSIVHNGAKF 15

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq 101002.*

1:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	15	17	AA97918
2	76	100.0	15	19	AA55765
3	76	100.0	460	16	AA69791
4	76	100.0	514	15	AA53690
5	76	100.0	514	16	AA74333
6	76	100.0	514	16	AA69792
7	76	100.0	514	17	AA93599
8	76	100.0	514	17	AA81586
9	76	100.0	514	20	AA25666
10	76	100.0	514	20	AA25667

11	72	94.7	14	18	AAW12539
12	72	94.7	14	23	AAU79882
13	72	94.7	134	17	AAW27371
14	63	82.9	514	18	AAW04346
15	63	82.9	514	19	AAW42122
16	61	80.3	453	22	AAW51693
17	61	80.3	507	22	AAW51691
18	59	77.6	20	19	AAW42187
19	54	71.1	15	17	AAW97919
20	50	65.8	15	17	AAW42186
21	49	64.5	20	19	AAW97917
22	43	56.6	415	21	AAW41325
23	43	56.6	431	20	AAW98178
24	43	56.6	431	21	AAW41324
25	43	56.6	431	21	ABW92704
26	43	56.6	463	21	AAW41323
27	42	55.3	1196	23	AAO14189
28	41	53.9	159	20	AAW35044
29	40	52.6	578	21	AAW95820
30	39	51.3	33	22	AAO77425
31	39	51.3	127	22	AAO7523
32	39	51.3	406	21	AAW21000
33	39	51.3	430	21	AAW20999
34	39	51.3	475	22	ABW59189
35	38	50.0	116	21	AAW75104
36	38	50.0	292	21	AAW52639
37	38	50.0	300	21	AAW52638
38	38	50.0	390	22	AAU33788
39	38	50.0	391	23	AAO20074
40	38	50.0	391	23	AAO20053
41	38	50.0	392	22	AAW82407
42	38	50.0	392	22	AAW82832
43	38	50.0	398	22	AAW37006
44	38	50.0	411	23	ABW58276
45	38	50.0	433	21	AAW51785

ALIGNMENTS

RESULT 1
AA97918
ID AA97918 standard; peptide; 15 AA.
XX
AC AA97918;
XX
DT 16-AUG-1996 (first entry)
XX
DE Japan cedar pollen mature allergen Cry j II amino acids 236-250.
XX
XX Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
KW Sugl pollinosis; diagnosis; treatment.
XX
OS Cryptomeria japonica.
XX
PN JP08047392-A.
XX
PD 20-FEB-1996.
XX
PF 07-NOV-1994; 94JP-0297840.
XX
PR 26-MAY-1994; 94JP-0134868.
XX
PR 05-NOV-1993; 93JP-0276773.
XX
PA (MEIP) MEIJI MILK PROD CO LTD.
XX
XX WPI; 1996-166249/17.
XX
XX Japan cedar pollen allergen Cry j II epitope - comprises at least
XX part of specified 460 amino acid protein
XX
XX Claim 8; Fig 4; 17pp; Japanese.

CC AAR97871-R97960 are overlapping peptides used for the epitope mapping
CC of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic
CC peptides of it are useful in the diagnosis, prevention and treatment
CC of Sugi pollinosis, the allergic reaction to Japan cedar pollen.
CC Significant regions of the allergen were identified using the
CC overlapping peptides of the full epitope derived from a Cry j II
CC antigen-specific T cell line. Amino acids 66-80 (AAR97884) and 186-200
CC (R978908) of the full mature 460 amino acid allergen are the most
CC allergenic of the 90 peptides tested.

XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 76; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. NO. 1.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAEVSIVHNGAKF 15
DB 1 SRAEVSIVHNGAKF 15

RESULT 2

AAW57765
ID AAW57765 standard; peptide; 15 AA.

XX
AC AAW57765;

XX
DT 17-SEP-1998 (first entry)

XX
DE Residues 236-250 of Cry j 2.

XX
KW Cry j 2; Japanese cedar pollen antigen; allergy; immunotherapy;

XX
KW HLA class II molecule.

XX
OS Cryptomeria japonica.

XX
PN WO9820902-A1.

XX
PD 22-MAY-1998.

XX
PF 12-NOV-1997; 97WO-JP04129.

XX
PR 13-NOV-1996; 96JP-0302053.

XX
PA (MEIP) MEIJI MILK PROD CO LTD.

XX
PI Dairiki K, Kino K, Kume A, Sone T;

XX
DR WPI; 1998-297617/26.

XX
PT Peptides derived from Japanese cedar pollen antigens are
PT immunotherapeutic agents - useful for allergy treatment and typing
PT HLA class II molecules in allergy sufferers

XX
PS Claim 12; Page 31; 50pp; Japanese.

XX
CC This sequence represents residues 236-250 of the Cry j 2 protein, and
CC is a peptide of the invention. The peptides are derived from Japanese
CC cedar pollen antigens, and are used as immunotherapeutic agents in the
CC treatment of allergy. The peptides can be used for identification and
CC typing of the particular HLA class II molecules in an allergy sufferer,
CC and also for peptide immunotherapy of an allergy. Using these peptides
CC the immunotherapy can be targeted more specifically to the requirements
CC of the individual patient, allowing more effective treatment of an
CC allergy, including those patients for whom treatment with a conventional
CC immunotherapeutic agent is ineffective.

XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 76; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. NO. 1.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAEVSIVHNGAKF 15
DB 1 SRAEVSIVHNGAKF 15

RESULT 3

AAW69791
ID AAR69791 standard; Protein; 460 AA.

XX
AC AAR69791;

XX
DT 27-SEP-1995 (first entry)

XX
DE Japonicum allergen residues 55-514.

XX
KW Japonicum allergen; residues 55-514; induced histamine release;

XX
KW antiallergic peptide; IgE cross-linking inhibition.

XX
OS Japonicum sp.

XX
PN WO9502412-A.

XX
PD 26-JAN-1995.

XX
PF 15-JUL-1994; 94WO-JP01164.

XX
PR 16-JUL-1993; 93JP-0177008.

XX
PR 01-SEP-1993; 93JP-0217725.

XX
PR 07-APR-1994; 94JP-0069336.

XX
PA (MEIP) MEIJI MILK PROD CO LTD.

XX
PI Kino K, Kohno Y, Komiyama N, Sone T;

XX
DR WPI; 1995-067159/09.

XX
DR N-PSDB; AAQ84044.

XX
PT Peptide antiallergic agent - inhibits cross-linking of allergen
PT with IgE antibody

XX
PS Disclosure; Pages 26-27; 46pp; Japanese.

XX
CC AAQ84044 encodes AAR69791 Japonicum allergen residues 55-514, from
CC which the antiallergic peptides AAR69845-R69809 were derived.
CC The peptides ability to inhibit the cross-linking of an allergen,
CC to an IgE antibody can be used in the prevention and treatment of
CC allergic diseases.

XX
SQ Sequence 460 AA;

Query Match 100.0%; Score 76; DB 16; Length 460;
Best Local Similarity 100.0%; Pred. NO. 7.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAEVSIVHNGAKF 15
DB 236 SRAEVSIVHNGAKF 250

RESULT 4

AAW53690
ID AAR53690 standard; Protein; 514 AA.

XX
AC AAR53690;

XX
DT 01-FEB-1995 (first entry)

XX
DE Japanese cedar pollen allergen Cry j II.

XX
KW Cedar pollinosis; diagnostic.

XX
OS Cryptomeria japonica.


```

PN MO9411512-A.
XX
PD 26-MAY-1994.
XX
PF 12-NOV-1993; 93WO-US11000.
XX
PR 12-NOV-1992; 92US-0975179.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Brauer A, Kuo M, Pollock J, Yeung S;
XX
DR WPI; 1994-183513/22.
XX
DR N-PSDB; AA066048.
XX
PT Allergic Cry j II protein and fragments from Japanese cedar
XX pollen - used to diagnose, treat and prevent Japanese cedar
XX pollinosis
XX
PS Claim 2; Fig 4; 89pp; English.
XX
CC The sequence is of a Japanese cedar pollen allergen Cry j
CC II. The protein and its fragments can be used for diagnosis and
CC treatment of Japanese cedar pollinosis and to identify similar
CC sequences in other plants.
CC See also AAR53692-6.
XX
SQ Sequence 514 AA;

Query Match 100.0%; Score 76; DB 15; Length 514;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 SRAEVSYYHNGAKF 15
Db 290 SRAEVSYYHNGAKF 304

RESULT 5
AAR74333
ID AAR74333 standard; Protein; 514 AA.
XX
AC AAR74333;
XX
DT 01-NOV-1995 (first entry)
XX
DE Japanese cedar pollen allergen.
XX
KW Japanese cedar; pollen; allergen; allergy; therapy; diagnostic;
XX deensulfier.
XX OS Cryptomeria japonica.
XX
PN EP655500-A.
XX
PD 31-MAY-1995.
XX
PF 03-NOV-1994; 94EP-0308117.
XX
PR 27-DEC-1993; 93JP-0346814.
XX PR 05-NOV-1993; 93JP-0239151.
XX PR 20-DEC-1993; 93JP-0344596.
XX
PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Kurimoto M, Namba M, Torigoe K;
XX
DR WPI; 1995-195588/26.
XX
DR N-PSDB; AAQ90156.
XX
PT New Japanese cedar pollen allergen polypeptide - and DNA coding for it,
XX useful for treatment and diagnosis of cedar pollen allergy
XX

```

```

PS Claim 5; Page 26-28; 41pp; English.
CC CC The gene encoding an allergen of Japanese cedar pollen was isolated
CC by PCR amplification using primers based on portions of the allergen
CC protein. The gene was used for recombinant allergen production in
CC E. coli (vector plasmid pKK-223-3).
XX
SQ Sequence 514 AA;

Query Match 100.0%; Score 76; DB 16; Length 514;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 SRAEVSYYHVGAKF 15
|||
Db 290 SRAEVSYYHVGAKF 304

RESULT 6
AAR69792
ID AAR69792 standard; Protein; 514 AA.
XX
AC AAR69792;
XX
DT 27-SEP-1995 (first entry)
XX
DE Japonicum allergen.
XX
KW Japonicum allergen; induced histamine release; antiallergic peptide;
XX IGE cross-linking inhibition.
XX Japonicum sp.
XX
OS WO9502412-A.
XX
PN 26-JAN-1995.
XX
PD 15-JUL-1994; 94WO-JP01164.
XX
PF 16-JUL-1993; 93JP-0177008.
XX
PR 01-SEP-1993; 93JP-0217725.
XX
PR 07-APR-1994; 94JP-0069336.
XX
PA (MEIP ) MEIJI MILK PROD CO LTD.
XX
XX Kino K, Kohno Y, Komiyama N, Sone T;
XX
XX WPI; 1995-067159/09.
XX
DR N-PSDB; AA084045; AA084046.
XX
XX Peptide antiallergic agent - inhibits cross-linking of allergen
XX with IGE antibody
XX
XX Example 3; Pages 27-28; 46pp; Japanese.
XX
PS AA084045 encodes AAR69792 japonicum allergen, from which the
CC antiallergic peptides AAR69845-R69809 were derived. The peptides ability
CC to inhibit the cross-linking of an allergen, to an IGE antibody can be
CC used in the prevention and treatment of allergic diseases.
XX
SQ Sequence 514 AA;

Query Match 100.0%; Score 76; DB 16; Length 514;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 SRAEVSYYHVGAKF 15
|||
Db 290 SRAEVSYYHVGAKF 304

RESULT 7
AAR935599

```

ID AAR93599 standard; Protein; 514 AA.
 XX AAR93599;
 AC
 XX 16-AUG-1996 (first entry)
 DT
 XX Japan cedar pollen Cry j II allergen.
 DE
 XX Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
 KM Sugi pollinosis; diagnosis; treatment.
 XX
 OS Cryptomeria japonica.
 XX
 FH Key Location/Qualifiers
 FT Peptide* 1..54
 FT /*label= sig_peptide
 FT 55..514
 FT Protein /*label= mat_protein
 FT
 XX JP08047392-A.
 PN
 XX 20-FEB-1996.
 PD
 XX 07-NOV-1994; 94JP-0297840.
 PF
 XX 26-MAY-1994; 94JP-0134868.
 PR
 XX 05-NOV-1993; 93JP-0276773.
 XX
 PA (MEIRI) MEIRI MILK PROD CO LTD.
 XX
 XX WPI; 1996-166249/17.
 DR N-PSDB; AAT18102.
 XX
 PT Japan cedar pollen allergen Cry j II epitope - comprises at least
 PT part of specified 460 amino acid protein
 XX
 PS Claim 1; Page 10-11; 17pp; Japanese.
 PS
 CC AAR93599 is a Japan cedar pollen Cry j II allergen which is useful
 CC in the diagnosis, prevention and treatment of Sugi pollinosis,
 CC the allergic reaction to Japan cedar pollen. Significant regions of
 CC the allergen were identified using overlapping peptides of the full
 CC epitope derived from a Cry j II antigen-specific T cell line
 CC (see AAR97871-R97860). Amino acids 66-80 (AAR97884) and 166-200
 CC (R978908) of the full mature 460 amino acid allergen are the most
 CC allergenic of the 90 peptides tested.
 CC
 XX
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 76; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 8e-06; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;
 QY 1 SRAEVSIVHNGAKF 15
 DB 290 SRAEVSIVHNGAKF 304
 RESULT 8
 ID AAR81586 standard; Protein; 514 AA.
 AC AAR81586;
 XX
 XX 24-MAY-1996 (first entry)
 DT
 XX Cedar pollen allergen A.
 DE
 XX Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
 KM antibody; pollinosis; therapy; immunotherapy.
 XX
 OS Cryptomeria japonica.
 XX

PN EP700929-A2.
 XX
 XX 13-MAR-1996.
 PD
 XX
 XX 08-SEP-1995; 95SEP-0306295.
 PF
 XX 14-JUL-1995; 95JP-0200221.
 PR 10-SEP-1994; 94JP-0242137.
 PR 14-JUL-1995; 95JP-0200204.
 XX
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 PI Hino K, Saito S, Taniguchi Y;
 DR WPI; 1996-140976/15.
 XX
 PT New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
 PT used for treating cedar pollinosis
 XX
 PS Claim 5; Page 29-30; 36pp; English.
 PS
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (AAR81586) and B (AAR81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
 CC for T-cell recognition, and homologous peptides (AAR81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 CC
 XX
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 76; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 8e-06; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;
 QY 1 SRAEVSIVHNGAKF 15
 DB 290 SRAEVSIVHNGAKF 304
 RESULT 9
 ID AAY25666 standard; protein; 514 AA.
 AC AAY25666;
 XX
 XX 30-SEP-1999 (first entry)
 DT
 XX Japanese cedar allergen 1076242 Cry j II precursor protein fragment.
 DE
 XX Major histocompatibility complex, class II; desensitising; human;
 KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX
 OS Cedrus sp.
 XX
 XX WO9934826-A1.
 PN
 XX 15-JUL-1999.
 PD
 XX 11-JAN-1999; 99WO-GB00080.
 PF
 XX 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PA Kay AB, Larche M;
 XX

XX DR WPI; 1999-458255/38.
 XX CC Desensitizing patients to polypeptide allergens
 PT XX Example 6; Page 74; 117pp; English.
 PS XX
 CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076242 Cry j II
 CC precursor.
 CC XX
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 76; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAEVSYPVHNGAKF 15
 DB 290 SRAEVSYPVHNGAKF 304
 RESULT 10
 AAY25667
 ID AAY25667 standard; protein; 514 AA.
 AC AAY25667;
 XX 30-SEP-1999 (first entry)
 DT 30-SEP-1999 (first entry)
 XX Japanese cedar allergen 1076241 Cry j II protein fragment.
 DE Japanese cedar allergen 1076241 Cry j II protein fragment.
 XX Major histocompatibility complex; class II; desensitizing; human;
 KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX Cedrus sp.
 OS Cedrus sp.
 XX WO9934826-A1.
 XX 15-JUL-1999.
 PD 15-JUL-1999.
 XX 11-JAN-1999; 99WO-GB00080.
 PF 11-JAN-1999; 99WO-GB00080.
 XX 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX Kay AB, Larche M;
 PI Kay AB, Larche M;
 XX WPI; 1999-458255/38.
 DR WPI; 1999-458255/38.
 XX Desensitizing patients to polypeptide allergens
 PT Example 6; Page 74; 117pp; English.
 PS

XX CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076241 Cry j II.
 CC XX
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 76; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAEVSYPVHNGAKF 15
 DB 290 SRAEVSYPVHNGAKF 304
 RESULT 11
 AAM12539
 ID AAM12539 standard; peptide; 14 AA.
 AC AAM12539;
 XX 30-APR-1997 (first entry)
 DT 30-APR-1997 (first entry)
 XX Japanese cedar pollen allergen-derived peptide 8.
 DE Japanese cedar pollen allergen-derived peptide 8.
 XX Japanese cedar; sugi; pollen; pollinosis; pollenosis; allergen;
 KW allergy; T-cell epitope; TCR; T-cell receptor; activation;
 KW immune tolerance; Cryj.
 XX Crytomeria japonica.
 OS Crytomeria japonica.
 XX JP08333391-A.
 PN JP08333391-A.
 XX 17-DEC-1996.
 PD 17-DEC-1996.
 XX 18-JUL-1995; 95JP-0181438.
 PF 18-JUL-1995; 95JP-0181438.
 XX 07-APR-1995; 95JP-0082519.
 PR 07-APR-1995; 95JP-0082519.
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX (SANY) SANKYO CO LTD.
 XX WPI; 1997-095487/09.
 DR WPI; 1997-095487/09.
 XX Peptide allergen derived from Japanese cedar pollen - causes T cell
 PT Response specific to cedar pollen, for treatment of pollenosis
 PS Claim 11; Page 19; 21pp; Japanese.
 XX The present sequence is one of 24 claimed peptides which were
 CC synthesised based on Japanese cedar pollen sequences. This peptide
 CC was shown to induce tolerance in a BALB/c mouse. The peptide
 CC produces little or no anaphylaxis.
 XX Sequence 14 AA;
 SQ Sequence 14 AA;
 Query Match 94.7%; Score 72; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 7.1e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAEVSYVHNGAKF 15
 |||||
 DB 1 RAEVSYVHNGAKF 14

RESULT 12

AAU79882 standard; Peptide; 14 AA.

AAU79882;

15-JUL-2002 (first entry)

Immune tolerance associated peptide #3.

Immune tolerance; transgenic plant; T-cell epitope;

allergic protein.

Cryptomeria japonica.

JP2002085068-A.

26-MAR-2002.

08-SEP-2000; 2000JP-0272863.

08-SEP-2000; 2000JP-0272863.

(SUMO) SUMITOMO CHEM CO LTD.

WPI; 2002-378273/41.

Induction of immune tolerance by oral administration of an effective

amount of transformed plant comprising a DNA coding for a polypeptide

having an amino acid sequence of T-cell epitope of an allergic protein

Example 4; Page 17; 19pp; Japanese.

The invention describes a method for induction of immune tolerance

in animals, including humans. The treatment involves by oral

administration of an effective amount of transformed plant comprising

a DNA coding for a polypeptide having an amino acid sequence of T-cell

epitope of an allergic protein. The introduced protein is not degraded

by digestive enzymes in the host. This sequence represents a peptide

associated with the method of inducing immune tolerance discussed in

the invention.

Sequence 14 AA;

Query Match 94.7%; Score 72; DB 23; Length 14;

Best Local Similarity 100.0%; Pred. No. 7.1e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAEVSYVHNGAKF 15
 |||||
 DB 1 RAEVSYVHNGAKF 14

24-MAR-1998 (first entry)

Multi-epitope peptide used as immunotherapeutic agent #3.

Multi-epitope peptide; immunotherapeutic agent; allergic disease;

T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN WO9732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

PA (MEIP) MEIJI MILK PROD CO LTD.

Dairiki K, Iwama A, Kino K, Kume A, Sone T;

WPI; 1997-470495/43.

Peptide immuno:therapeutic agent to treat allergic diseases -

contains multi-epitope peptide containing T cell epitope regions

from different allergens

Claim 6; Page 32; 58pp; Japanese.

The present sequence represents a multi-epitope peptide which is used as

a new immunotherapeutic agent. It comprises T cell epitope regions from 2

or more different allergens (preferably linked via arginine or lysine

dimers), where the T cell epitope regions have a positivity index

greater than 100 as measured in a patient group responding to the

allergen; have at least 70% reactivity with lymphocytes from patients

responding to the allergen; and are not reactive with immunoglobulin E

(IgE) antibodies from patients responsive to the allergen. The agent can

be used to prevent and treat a wide variety of allergic diseases, e.g. by

desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

Sequence 134 AA;

Query Match 94.7%; Score 72; DB 18; Length 134;

Best Local Similarity 100.0%; Pred. No. 9.7e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RAEVSYVHNGAKF 15
 |||||
 DB 103 RAEVSYVHNGAKF 116

RESULT 14

AAW04346 standard; Protein; 514 AA.

AAW04346;

28-NOV-1996 (first entry)

Chamaecyparis obtusa pollen allergen Cha o II.

Pollen allergen; Cha o II; T-cell epitope; prevention; treatment;

pollinosis.

Chamaecyparis obtusa.

JP08176192-A.

09-JUL-1996.

21-DEC-1994; 94JP-0335089.

21-DEC-1994; 94JP-0335089.

(MEIP) MEIJI MILK PROD CO LTD.

WPI; 1996-368225/37.

N-PSDB; AAT38521.

PT DNA encoding chamaecyparis obtusa pollen allergen - T cell
 PT epitope(s) of which are useful in development of preventative and
 PT treating agent for C. obtusa pollen pollinosis

PS Claim 14; Pages 14-15; 17pp; Japanese.

CC The present sequence is the C. obtusa pollen allergen
 CC Cha o II, the T-cell epitopes of which can be used in the
 CC development of a preventative and treating agent for C. obtusa
 CC pollen pollinosis. C. obtusa pollen (2.4 kg) was degraded with
 CC diethyl ether, and dried at room temp. overnight. Cha o II was
 CC sepd. from it and purified. RNA was extracted from C. obtusa
 CC pollen, and mRNA and cDNA derived.

SQ Sequence 514 AA;

Query Match 82.9%; Score 63; DB 19; Length 514;

Best Local Similarity 86.7%; Pred. No. 0.0023;

Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SRAEVSIVHNGAKF 15

DB 290 SRAEVSIVHNGAKF 304

RESULT 15

AAW42122

ID AAW42122 standard; protein; 514 AA.

AC AAW42122;

DT 16-JUN-1998 (first entry)

DE Japanese cypress pollen antigen Cha2.

KW Japanese cypress pollen; antigen; T-cell epitope; Cha2; Cha2;
 KW diagnosis; allergy; spring tree pollen disease; pollinosis.

OS Chamaecyparis obtusa.

PN WO9747648-A1.

PD 18-DEC-1997.

PF 12-JUN-1997; 97WO-JP02031.

PK 14-JUN-1996; 96JP-0153527.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Kino K;

DR WPI, 1998-052242/05.

PT T-cell epitope peptide portion of Japanese cypress pollen antigens
 PT Cha2 and Cha2 - used for diagnosis and treatment of spring tree
 PT pollen disease

PS Example 1; Page 15-18; 71pp; Japanese.

CC The present sequence represents Japanese cypress pollen antigen Cha2.
 CC The present invention describes peptides which correspond to the T-cell
 CC epitope sites on Japanese cypress pollen antigens Cha2 and Cha2. The
 CC peptides can be used as a reagent for the diagnosis of allergy to
 CC Japanese cypress pollen, and as an antigen in the treatment and
 CC prevention of spring tree pollen disease in which the pollinosis
 CC involves reactivity to Japanese cypress pollen.

SQ Sequence 514 AA;

Query Match 82.9%; Score 63; DB 19; Length 514;

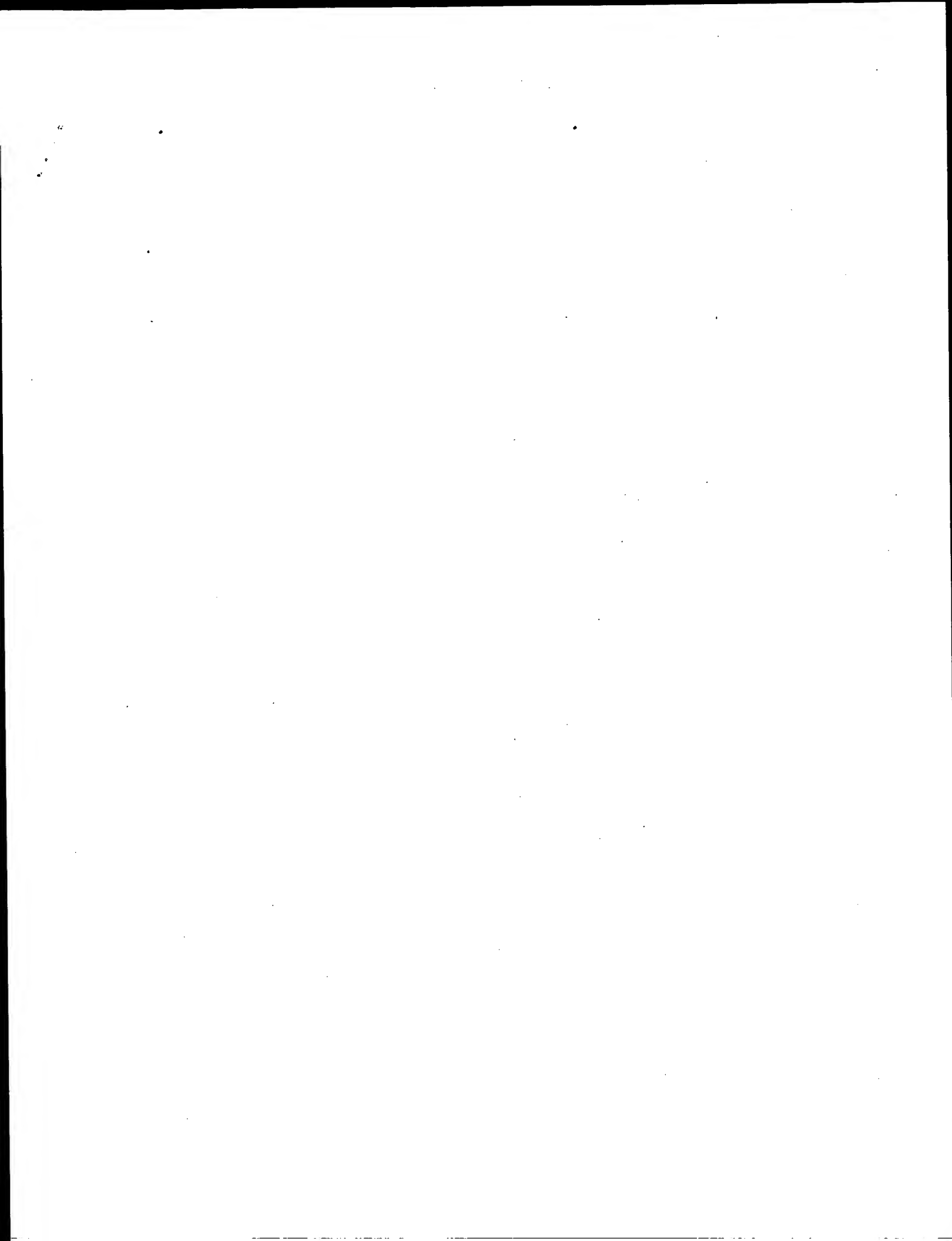
Best Local Similarity 86.7%; Pred. No. 0.0023;

Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SRAEVSIVHNGAKF 15

DB 290 SRAEVSIVHNGAKF 304

Search completed: April 20, 2003, 13:06:12
 Job time: 19.1974 secs



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds

(Without alignments)
109,838 Million cell updates/sec

Title: US-09-142-524D-152

Perfect score: 67

Sequence: 1 LSDSLKLTGSKIAS 15

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database : *
1: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDs2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	15	17	AAAR97939
2	67	100.0	15	19	AAAR57770
3	67	100.0	460	16	AAAR69791
4	67	100.0	514	15	AAAR53690
5	67	100.0	514	16	AAAR74333
6	67	100.0	514	16	AAAR69792
7	67	100.0	514	17	AAAR93599
8	67	100.0	514	17	AAAR1586
9	67	100.0	514	20	AAI25667
10	65	97.0	514	20	AAI25666

11	59	88.1	17	17	AAAR1583	Cedar pollen aller
12	54	80.6	20	19	AAAR42197	T-cell epitope pep
13	54	80.6	514	17	AAAR04346	Chamaecyparis obtu
14	54	80.6	514	19	AAAR42122	Japanese cypress p
15	53	79.1	17	13	AAAR80348	Sugi allergen prot
16	53	79.1	80	18	AAAR27369	Multi-epitope pep
17	51	76.1	453	22	AAAR5693	Juniperus ashei N-
18	51	76.1	507	22	AAAR5691	Juniperus ashei Ju
19	49	73.1	11	17	AAAR1577	Cedar pollen aller
20	49	73.1	17	17	AAAR1594	Cedar pollen aller
21	45	67.2	12	18	AAAR14301	Japanese cedar pol
22	45	67.2	15	17	AAAR97940	Japan cedar pollen
23	44	65.7	15	17	AAAR97938	Japan cedar pollen
24	44	65.7	15	19	AAAR57768	Residues 336-350 o
25	44	65.7	35	19	AAAR80342	Sugi allergen prot
26	44	65.7	35	19	AAAR80344	Sugi allergen prot
27	44	65.7	152	19	AAAR8539	S. pneumoniae sul
28	44	65.7	242	18	AAAR11194	Streptococcus pneu
29	44	65.7	242	21	AAAR11732	Staphylococcus aur
30	44	65.7	299	22	AAAR37365	Sugi allergen prot
31	44	65.7	299	22	AAAR37365	Sugi allergen prot
32	43	64.2	47	19	AAAR80356	Japanese cedar pol
33	43	64.2	11	18	AAAR14302	Japanese cedar pol
34	41	61.2	12	18	AAAR14303	Cryptomeria japon
35	41	61.2	12	22	AAAR69109	Sugi allergen prot
36	41	61.2	33	19	AAAR80340	Sugi allergen prot
37	41	61.2	33	19	AAAR80339	Sugi allergen prot
38	41	61.2	33	19	AAAR80341	Sugi allergen prot
39	41	61.2	35	19	AAAR80343	Sugi allergen prot
40	41	61.2	47	19	AAAR80350	Sugi allergen prot
41	41	61.2	47	19	AAAR80352	Sugi allergen prot
42	41	61.2	47	19	AAAR80352	Sugi allergen prot
43	41	61.2	47	19	AAAR80353	Sugi allergen prot
44	41	61.2	47	19	AAAR80357	Sugi allergen prot
45	41	61.2	51	21	AAAR3877	Artificial sequenc

ALIGNMENTS

```
RESULT 1
AAAR97939
ID AAAR97939 standard; peptide; 15 AA.
XX
AC AAAR97939;
XX
DT 16-AUG-1996 (first entry)
XX
DE Japan cedar pollen mature allergen Cry j II amino acids 341-355.
XX
KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
XX Sugi pollinosis; diagnosis; treatment.
XX
OS Cryptomeria japonica.
XX
PN JP08047392-A.
XX
PD 20-FEB-1996.
XX
PF 07-NOV-1994; 94JP-0297840.
XX
PR 26-MAY-1994; 94JP-0134868.
XX
PR 05-NOV-1993; 93JP-0276773.
XX
PA (MEIP) MEIJI MILK PROD CO LTD.
XX
DR WPI, 1996-166249/17.
XX
PT Japan cedar pollen allergen Cry j II epitope - comprises at least
XX part of specified 460 amino acid protein
XX
PS Claim 8; Fig 5; 17pp; Japanese.
XX
```

CC AAR97871-R97860 are overlapping peptides used for the epitope mapping
 CC of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic
 CC peptides of it are useful in the diagnosis, prevention and treatment
 CC of Sugi pollinosis, the allergic reaction to Japan cedar pollen.
 CC Significant regions of the allergen were identified using the
 CC overlapping peptides of the full epitope derived from a Cry j II
 CC antigen-specific T cell line. Amino acids 66-80 (AAR97884) and 186-200
 CC (R978908) of the full mature 460 amino acid allergen are the most
 CC allergenic of the 90 peptides tested.

CC Sequence 15 AA;

Query Match 100.0%; Score 67; DB 17; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.5e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LSDISLKTSGKIAS 15
 |||||
 Db 1 LSDISLKTSGKIAS 15

RESULT 2

ID AAW57770 standard; peptide; 15 AA.

AC AAW57770;

DT 17-SEP-1998 (first entry)

DE Residues 341-355 of Cry j 2.

XX Cry j 2; Japanese cedar pollen antigen; allergy; immunotherapy;

KW HLA class II molecule.

OS Cryptomeria japonica.

FN WO9820902-A1.

PD 22-MAY-1998.

PF 12-NOV-1997; 97WO-JP04129.

PR 13-NOV-1996; 96JP-0302053.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Kino K, Kume A, Sone T;

DR WPI, 1998-297617/26.

PT Peptides derived from Japanese cedar pollen antigens are
 PT immunotherapeutic agents - useful for allergy treatment and typing
 PT HLA class II molecules in allergy sufferers

PS Claim 5; Page 33; 50pp; Japanese.

XX This sequence represents residues 341-355 of the Cry j 2 protein, and
 CC is a peptide of the invention. The peptides are derived from Japanese
 CC cedar pollen antigens, and are used as immunotherapeutic agents in the
 CC treatment of allergy. The peptides can be used for identification and
 CC typing of the particular HLA class II molecules in an allergy sufferer,
 CC and also for peptide immunotherapy of an allergy. Using these peptides
 CC the immunotherapy can be targeted more specifically to the requirements
 CC of the individual patient, allowing more effective treatment of an
 CC allergy, including those patients for whom treatment with a conventional
 CC immunotherapeutic agent is ineffective.

SO Sequence 15 AA;

Query Match 100.0%; Score 67; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LSDISLKTSGKIAS 15
 |||||
 Db 1 LSDISLKTSGKIAS 15

RESULT 3

ID AAR69791 standard; Protein; 460 AA.

AC AAR69791;

DT 27-SEP-1995 (first entry)

DE Japonicum allergen residues 55-514.

XX Japonicum allergen; residues 55-514; induced histamine release;

KW antiallergic peptide; IgE cross-linking inhibition.

OS Japonicum sp.

PN WO9502412-A.

PD 26-JAN-1995.

PF 15-JUL-1994; 94WO-JP01164.

PR 16-JUL-1993; 93JP-0177008.

PR 01-SEP-1993; 93JP-0217725.

PR 07-APR-1994; 94JP-0069336.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Kino K, Kohno Y, Komiyama N, Sone T;

DR WPI, 1995-067159/09.

DR N-PSDB; AAO84044.

XX Peptide antiallergic agent - inhibits cross-linking of allergen

PT with IgE antibody

PS Disclosure; Pages 26-27; 46pp; Japanese.

XX AAO84044 encodes AAR69791 Japonicum allergen residues 55-514, from

CC which the antiallergic peptides AAR69845-R69809 were derived.

CC The peptides ability to inhibit the cross-linking of an allergen,

CC to an IgE antibody can be used in the prevention and treatment of

CC allergic diseases.

SO Sequence 460 AA;

Query Match 100.0%; Score 67; DB 16; Length 460;
 Best Local Similarity 100.0%; Pred. No. 0.0007;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LSDISLKTSGKIAS 15
 |||||
 Db 341 LSDISLKTSGKIAS 355

RESULT 4

AAAR53690 standard; Protein; 514 AA.

AC AAR53690;

DT 01-FEB-1995 (first entry)

DE Japanese cedar pollen allergen Cry j II.

XX Cedar pollinosis; diagnostic.

OS Cryptomeria japonica.

PN WO9411512-A.
 XX 26-MAY-1994.
 XX
 XX
 PF 12-NOV-1993; 93WO-US11000.
 XX
 PR 12-NOV-1992; 92US-0975179.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Brauer A, Kuo M, Pollock J, Yeung S;
 XX
 DR WPI; 1994-183513/22.
 DR N-PSDB; AAQ66048.
 XX
 PT Allergenic Cry j II protein and fragments from Japanese cedar
 PT pollen - used to diagnose, treat and prevent Japanese cedar
 XX pollinosis
 XX
 PS Claim 2; Fig 4; 89pp; English.
 XX
 CC The sequence is of a Japanese cedar pollen allergen Cry j
 CC II. The protein and its fragments can be used for diagnosis and
 CC treatment of Japanese cedar pollinosis and to identify similar
 CC sequences in other plants.
 CC See also AAR53692-6.
 XX
 SQ Sequence 514 AA;
 XX
 Query Match 100.0%; Score 67; DB 15; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.00079;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LSDISLKTSGKIAS 15
 DB 395 LSDISLKTSGKIAS 409
 XX
 RESULT 5
 AAR74333
 ID AAR74333 standard; Protein; 514 AA.
 XX
 AC AAR74333;
 XX
 DT 01-NOV-1995 (first entry)
 XX
 DE Japanese cedar pollen allergen.
 XX
 KW Japanese cedar; pollen; allergen; allergy; therapy; diagnostic;
 KW desensitizer.
 XX
 OS Cryptomeria japonica.
 XX
 PN EP65500-A.
 XX
 PD 31-MAY-1995.
 XX
 PF 03-NOV-1994; 94EP-0308117.
 XX
 PR 27-DEC-1993; 93JP-0346814.
 PR 05-NOV-1993; 93JP-0299151.
 PR 20-DEC-1993; 93JP-0344596.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 PI Kurimoto M, Namba M, Torigoe K;
 XX
 DR WPI; 1995-195588/26.
 DR N-PSDB; AAQ90156.
 XX
 PT New Japanese cedar pollen allergen polypeptide - and DNA coding for it,
 PT useful for treatment and diagnosis of cedar pollen allergy
 XX

PS Claim 5; Page 26-28; 41pp; English.
 XX
 CC The gene encoding an allergen of Japanese cedar pollen was isolated
 CC by PCR amplification using primers based on portions of the allergen
 CC protein. The gene was used for recombinant allergen production in
 CC E. coli (vector plasmid pKK-223-3).
 XX
 SQ Sequence 514 AA;
 XX
 Query Match 100.0%; Score 67; DB 16; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.00079;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LSDISLKTSGKIAS 15
 DB 395 LSDISLKTSGKIAS 409
 XX
 RESULT 6
 AAR69792
 ID AAR69792 standard; Protein; 514 AA.
 XX
 AC AAR69792;
 XX
 DT 27-SEP-1995 (first entry)
 XX
 DE Japonicum allergen.
 XX
 KW Japonicum allergen; induced histamine release; antiallergic peptide;
 KW Ige cross-linking inhibition.
 XX
 OS Japonicum sp.
 XX
 PN WO9502412-A.
 XX
 PD 26-JAN-1995.
 XX
 PF 15-JUL-1994; 94WO-JP01164.
 XX
 PR 16-JUL-1993; 93JP-0177008.
 PR 01-SEP-1993; 93JP-0217725.
 PR 07-APR-1994; 94JP-0069336.
 XX
 PA (MEIP) MEIJI MILK PROD CO LTD.
 XX
 PI Kino K, Kohno Y, Komiyama N, Sone T;
 XX
 DR WPI; 1995-067159/09.
 DR N-PSDB; AAQ84045; AAQ84046.
 XX
 PT Peptide antiallergic agent - inhibits cross-linking of allergen
 PT with Ige antibody
 XX
 PS Example 3; Pages 27-28; 46pp; Japanese.
 XX
 CC AAQ84045 encodes AAR69792 Japonicum allergen, from which the
 CC antiallergic peptides AAR69845-R69809 were derived. The peptides ability
 CC to inhibit the cross-linking of an allergen, to an Ige antibody can be
 CC used in the prevention and treatment of allergic diseases.
 XX
 SQ Sequence 514 AA;
 XX
 Query Match 100.0%; Score 67; DB 16; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.00079;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LSDISLKTSGKIAS 15
 DB 395 LSDISLKTSGKIAS 409
 XX
 RESULT 7
 AAR93599

ID AAR93599 standard; Protein; 514 AA.
 XX AAR93599;
 AC
 XX 16-AUG-1996 (first entry)
 DT
 XX Japan cedar pollen Cry j II allergen.
 DE
 XX Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
 KM Sugi pollinosis; diagnosis; treatment.
 XX
 OS Cryptomeria japonica.
 XX
 FT Key Location/Qualifiers
 FT Peptide 1..54
 FT /*label= sig_peptide
 FT Protein 55..514
 FT /*label= mat_protein
 XX
 PN JP08047392-A.
 XX
 PD 20-FEB-1996.
 XX
 PF 07-NOV-1994; 94JP-0297840.
 XX
 PR 26-MAY-1994; 94JP-0134868.
 PR 05-NOV-1993; 93JP-0276773.
 XX
 PA (MEIP) MEIJI MILK PROD CO LTD.
 XX
 DR WPI; 1996-166249/17.
 DR N-PSDB; AAT18102.
 XX
 PT Japan cedar pollen allergen Cry j II epitope - comprises at least
 PT part of specified 460 amino acid protein
 XX
 PS Claim 1; Page 10-11; 17pp; Japanese.
 XX
 CC AAR93599 is a Japan cedar pollen Cry j II allergen which is useful
 CC in the diagnosis, prevention and treatment of Sugi pollinosis,
 CC the allergic reaction to Japan cedar pollen. Significant regions of
 CC the allergen were identified using overlapping peptides of the full
 CC epitope derived from a Cry j II antigen-specific T cell line
 CC (see AAR97871-R97860). Amino acids 66-80 (AAR97884) and 186-200
 CC (R978908) of the full mature 460 amino acid allergen are the most
 CC allergenic of the 90 peptides tested.
 CC
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 67; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.00079; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;
 OY 1 LSDISLKLTSKGIAS 15
 DB 395 LSDISLKLTSKGIAS 409
 RESULT 8
 ID AAR81586 standard; Protein; 514 AA.
 XX AAR81586;
 AC
 XX 24-MAY-1996 (first entry)
 DT
 XX Cedar pollen allergen A.
 DE
 XX Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
 KM antibody; pollinosis; therapy; immunotherapy.
 XX
 OS Cryptomeria japonica.
 XX

PN EP700929-A2.
 XX
 PD 13-MAR-1996.
 XX
 PF 08-SEP-1995; 95EP-0306295.
 XX
 PR 14-JUL-1995; 95JP-0200221.
 PR 10-SEP-1994; 94JP-0242137.
 PR 14-JUL-1995; 95JP-0200204.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 PI Hino K, Saito S, Taniguchi Y;
 XX
 DR WPI; 1996-140976/15.
 XX
 PT New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
 PT used for treating cedar pollinosis
 XX
 PS Claim 5; Page 29-30; 36pp; English.
 XX
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (AAR81586) and B (AAR81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequence (AAR81573-79) essential
 CC for T-cell recognition, and homologous peptides (AAR81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 CC
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 67; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.00079; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;
 OY 1 LSDISLKLTSKGIAS 15
 DB 395 LSDISLKLTSKGIAS 409
 RESULT 9
 ID AAY25667 standard; Protein; 514 AA.
 XX AAY25667;
 AC
 XX 30-SEP-1999 (first entry)
 DT
 XX Japanese cedar allergen 1076241 Cry j II protein fragment.
 DE
 XX Major histocompatibility complex; class II; desensitizing; human;
 KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX
 OS Cedrus sp.
 XX
 PN WO934826-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 11-JAN-1999; 99WO-GB00080.
 XX
 PR 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Kay AB, Larche M;

XX WPI; 1999-458255/38.
 XX Desensitizing patients to polypeptide allergens
 PT
 XX
 PS Example 6; Page 74; 117pp; English.
 XX

CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076242 Cry j II.
 CC
 SQ Sequence 514 AA;

Query Match 100.0%; Score 67; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.00079;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSDISLKTSGKIAS 15
 |||||:|||||
 Db 395 LSDISLKTSGKIAS 409

RESULT 10
 AAY25666
 ID AAY25666 standard; protein; 514 AA.
 XX
 AC AAY25666;

DT 30-SEP-1999 (first entry)
 XX

DE Japanese cedar allergen 1076242 Cry j II precursor protein fragment.
 XX

KM Major histocompatibility complex; class II; desensitizing; human;
 KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 KM
 OS Cedrus sp.
 XX
 PN WO9934826-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 11-JAN-1999; 99WO-GB00080.
 XX
 PR 21-SEP-1998; 98GB-0020474.
 XX
 PR 09-JAN-1998; 98GB-0000445.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Kay AB, Larche M;
 XX
 DR WPI; 1999-458255/38.
 XX
 PT Desensitizing patients to polypeptide allergens
 XX
 PS Example 6; Page 74; 117pp; English.
 XX

CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076242 Cry j II
 CC precursor.
 CC
 SQ Sequence 514 AA;

Query Match 97.0%; Score 65; DB 20; Length 514;
 Best Local Similarity 93.3%; Pred. No. 0.0018;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSDISLKTSGKIAS 15
 |||||:|||||
 Db 395 LSDISLKTSGKIAS 409

RESULT 11
 AAR81583
 ID AAR81583 standard; Peptide; 17 AA.
 XX
 AC AAR81583;

DT 24-MAY-1996 (first entry)
 XX

DE Cedar pollen allergen peptide 11 (T-cell epitope).
 XX

KM Cedar; Cryptomeria japonica; pollen; allergen; immunoglobulin E;
 KM IgE; T-cell epitope; antibody; pollinosis; therapy; immunotherapy.
 XX
 OS Synthetic.
 OS
 PN EP700929-A2.
 XX
 PD 13-MAR-1996.
 XX
 PF 08-SEP-1995; 95EP-0306295.
 XX
 PR 14-JUL-1995; 95JP-0200221.
 XX
 PR 10-SEP-1994; 94JP-0242137.
 XX
 PR 14-JUL-1995; 95JP-0200204.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 PI Hino K, Saito S, Taniguchi Y;
 XX
 DR WPI; 1996-140976/15.
 XX
 PT New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
 PT used for treating cedar pollinosis
 XX
 PS Claim 4; Page 29; 36pp; English.
 XX
 DR Synthetic peptides based on portions of cedar pollen allergens A
 CC (AAR81586) and B (AAR81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
 CC for T-cell recognition, and homologous peptides (AAR81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar

CC pollinosis, avoiding side-effects such as anaphylaxis.

XX Sequence 17 AA;

Query Match 88.1%; Score 59; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.0049;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DISLKTSGKIAS 15

DB 1 DISLKTSGKIAS 13

RESULT 12

AAW42197
ID AAW42197 standard; peptide; 20 AA.

AC AAW42197;

DT 16-JUN-1998 (first entry)

XX T-cell epitope peptide 77 from Japanese cypress pollen antigen Cha02.

KW Japanese cypress pollen; antigen; T-cell epitope; Cha01; Cha02;

KW diagnosis; allergy; spring tree pollen disease; pollinosis.

OS Chamaecyparis obtusa.

XX WO9747648-A1.

XX 18-DEC-1997.

XX 12-JUN-1997; 97WO-JP02031.

XX 14-JUN-1996; 96JP-0153527.

XX (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Kino K;

DR WPI; 1998-052242/05.

PT T-cell epitope peptide portion of Japanese cypress pollen antigens

PT Cha01 and Cha02 - used for diagnosis and treatment of spring tree

PS pollen disease

XX Claim 2; Page 50; 71pp; Japanese.

CC The present sequence represents a T-cell epitope peptide from Japanese
CC cypress pollen antigen Cha02. The present invention describes peptides
CC which correspond to the T-cell epitope sites on Japanese cypress pollen
CC antigens Cha01 and Cha02. The peptides can be used as a reagent for the
CC diagnosis of allergy to Japanese cypress pollen, and as an antigen in
CC the treatment and prevention of spring tree pollen disease in which the
CC pollinosis involves reactivity to Japanese cypress pollen.

XX Sequence 20 AA;

Query Match 80.6%; Score 54; DB 19; Length 20;

Best Local Similarity 80.0%; Pred. No. 0.0049;

Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSDSLKTSGKIAS 15

DB 5 LSNVSLKTSGKPAS 19

RESULT 13

AAW04346
ID AAW04346 standard; Protein; 514 AA.

AC AAW04346;

XX

DT 28-NOV-1996 (first entry)

XX Chamaecyparis obtusa pollen allergen Cha o II.

XX Pollen allergen; Cha o II; T-cell epitope; prevention; treatment;

XX pollinosis.

OS Chamaecyparis obtusa.

XX JP08176192-A.

XX 09-JUL-1996.

XX 21-DEC-1994; 94JP-0335089.

XX 21-DEC-1994; 94JP-0335089.

XX (MEIP) MEIJI MILK PROD CO LTD.

XX WPI; 1996-368225/37.

XX N-PSDB; AAT38521.

PT DNA encoding chamaecyparis obtusa pollen allergen - T cell

PT epitope(s) of which are useful in development of preventative and

PS creating agent for C. obtusa pollen pollinosis

XX Claim 14; Pages 14-15; 17pp; Japanese.

CC The present sequence is the C. obtusa pollen allergen
CC Cha o II, the T-cell epitopes of which can be used in the
CC development of a preventative and treating agent for C. obtusa
CC pollen pollinosis. C. obtusa pollen (2.4 Kg) was degassed with
CC diethyl ether, and dried at room temp. overnight. Cha o II was
CC sepd. from it and purified. RNA was extracted from C. obtusa
CC pollen, and mRNA and cDNA derived.

XX Sequence 514 AA;

Query Match 80.6%; Score 54; DB 17; Length 514;

Best Local Similarity 80.0%; Pred. No. 0.19;

Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSDSLKTSGKIAS 15

DB 395 LSNVSLKTSGKPAS 409

RESULT 14

AAW42122
ID AAW42122 standard; protein; 514 AA.

AC AAW42122;

DT 16-JUN-1998 (first entry)

XX Japanese cypress pollen antigen Cha02.

XX Japanese cypress pollen; antigen; T-cell epitope; Cha01; Cha02;

XX diagnosis; allergy; spring tree pollen disease; pollinosis.

OS Chamaecyparis obtusa.

XX WO9747648-A1.

XX 18-DEC-1997.

XX 12-JUN-1997; 97WO-JP02031.

XX 14-JUN-1996; 96JP-0153527.

XX (MEIP) MEIJI MILK PROD CO LTD.

XX Dairiki K, Kino K;

PI

XX WPI; 1998-052242/05.
 DR T-cell epitope peptide portion of Japanese cypress pollen antigens
 XX Chaol and Chaol2 - used for diagnosis and treatment of spring tree
 PT pollen disease
 PT pollen disease
 XX Example 1; Page 15-18, 71pp; Japanese.
 PS
 CC The present sequence represents Japanese cypress pollen antigen Chaol2.
 CC The present invention describes peptides which correspond to the T-cell
 CC epitope sites on Japanese cypress pollen antigens Chaol and Chaol2. The
 CC peptides can be used as a reagent for the diagnosis of allergy to
 CC Japanese cypress pollen, and as an antigen in the treatment and
 CC prevention of spring tree pollen disease in which the pollinosis
 CC involves reactivity to Japanese cypress pollen.
 XX
 SQ Sequence 514 AA;

Query Match Best Local Similarity 80.6%; Score 54; DB 19; Length 514;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LSDISLKTSGKIAS 15
 ||:|||||
 DB 395 LSNVSLKTSGKPDAS 409

RESULT 15

AAW80348
 ID AAW80348 standard; peptide; 17 AA.

AC AAW80348;

DT 11-JAN-1999 (first entry)

DE Sugi allergen protein Cryj2 derived epitope for T cells.

XX T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;
 KW sugi-pollinosis; allergic reaction; pollen.

OS Synthetic.

PN JP10259198-A.

PD 29-SEP-1998.

PF 22-DEC-1997; 97JP-0353448.

PR 24-DEC-1996; 96JP-0343441.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 (SANY) SANKYO CO LTD.

DR WPI; 1998-577037/49.

PT A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollinosis

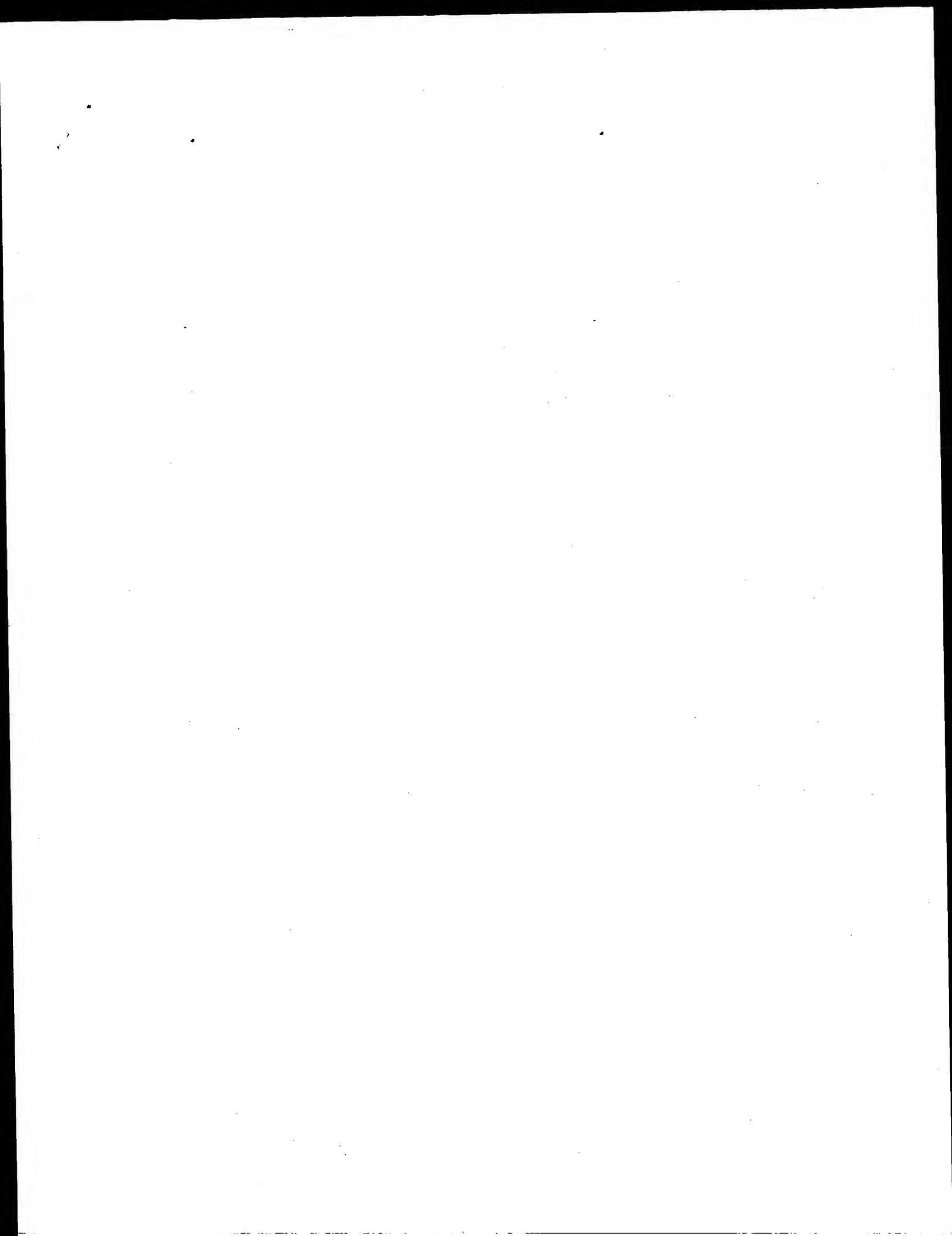
PS Claim 7; Page 18; 21pp; Japanese.

CC AAW80339-58 represent epitopes for T cells, derived from the sugi
 CC allergen proteins Cryj1 (AAW80339-44, AAW80350-53 and AAW80356-58) and
 CC Cryj2 (AAW80345-49 and AAW80354-55). The peptides are useful for the
 CC treatment of sugi-pollinosis, an allergic reaction of the body to
 CC pollen.

SQ Sequence 17 AA;

QY 4 ISLKTSGKIAS 15
 |||||||
 DB 1 ISLKTSGKIAS 12
 Search completed: April 20, 2003, 13:06:13
 Job time : 19.1974 secs

Query Match Best Local Similarity 79.1%; Score 53; DB 19; Length 17;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds
(without alignments) 237.215 Million cell updates/sec

Title: US-09-142-524D-28
Perfect score: 81
Sequence: 1 RPLMIIFSGNNMIKL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	374	2	JC2124
2	81	100.0	374	2	JC2123
3	58	71.6	397	2	S26211
4	57	70.4	542	2	T06728
5	56	69.1	455	2	T00856
6	56	69.1	463	2	T46165
7	55	67.9	449	2	S27098
8	54	66.7	398	2	B53240
9	54	66.7	398	2	B39099
10	53	65.4	392	2	D53240
11	53	65.4	472	2	T51456
12	52	64.2	396	2	A39099
13	52	64.2	397	2	C53240
14	52	64.2	397	2	C38099
15	52	64.2	434	2	S29612
16	51	63.0	438	2	S43335
17	51	63.0	459	2	G86278
18	50	61.7	274	2	T03932
19	50	61.7	431	2	T86179
20	49	60.5	404	2	T05556
21	49	60.5	501	2	T34741
22	48	59.3	397	2	E53240
23	47	58.0	385	2	A91006
24	47	58.0	385	2	B85850
25	46	56.8	385	2	A64981
26	46	56.8	398	2	T07058
27	45	55.6	404	2	S12209
28	45	55.6	251	2	C86248
29	45	55.6	374	2	T05240

ALIGNMENTS

30	45	55.6	374	2	H85148	probable pectate 1
31	44	54.3	450	2	T09524	probable pectate 1
32	43	53.1	341	2	T47653	pectate lyase-like
33	43	53.1	418	2	T07701	pectate lyase (EC
34	43	53.1	839	2	T21207	hypothetical prote
35	42	51.9	344	2	D86141	protein T25K16.7
36	42	51.9	350	2	D84812	probable peroxidase
37	42	51.9	394	2	T49115	pectate lyase like
38	42	51.9	394	2	T49116	pectate lyase like
39	42	51.9	550	2	T40486	phosphotriesterid-d
40	41	50.6	137	2	B75075	hypothetical prote
41	41	50.6	174	2	A86693	hypothetical prote
42	41	50.6	174	2	A86427	probable pectate 1
43	41	50.6	368	2	G86253	hypothetical prote
44	41	50.6	390	2	H86253	caton transport A
45	41	50.6	915	2	H82104	vacuolar protein s
			1466	2	T39557	

RESULT 1

JC2124
major allergen Cry j I precursor (clone PCCI-15) - Japanese cedar
C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000
C/Accession: JC2124
R/Source, T.: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 199, 619-625, 1994
A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japan
A/Reference number: JC2123; MUID:94183234; PMID:8135802
A/Accession: JC2124
A/Molecule type: mRNA
A/Residues: 1-374 <SON>
A/Cross-References: GB:D26545; NID:G493633; PIDN:BA05543.1; PID:G493634
A/Experimental source: pollen
A/Note: the authors described carbohydrate binding site for residue 279
C/Superfamily: pectate lyase LAM59
C/Keywords: glycoprotein; pollen
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-374/Product: major allergen Cry j I (clone PCCI-15) #status predicted <MAT>
F:158/191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. NO. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPLMIIFSGNNMIKL 15
DB 87 RPLMIIFSGNNMIKL 101

RESULT 2

JC2123
major allergen Cry j I precursor (clone PCCI-2-2) - Japanese cedar
C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000
C/Accession: JC2123; PC2065
R/Source, T.: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 199, 619-625, 1994
A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japan
A/Reference number: JC2123; MUID:94183234; PMID:8135802
A/Accession: JC2123
A/Molecule type: mRNA
A/Residues: 1-374 <SON>
A/Cross-References: GB:D26544; NID:G493631; PIDN:BA05542.1; PID:G493632
A/Experimental source: pollen
A/Accession: PC2065
A/Molecule type: protein
A/Residues: 22-53; 58-81; 219-232; 236-258; 299-307; 346-372 <SO2>
A/Note: the authors described carbohydrate binding site for residue 279
C/Superfamily: pectate lyase LAM59
C/Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-37/Product: major allergen Cry I (clone PCCI-2-2) #status predicted <MA>
 F:150,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 374;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPLWIFSGNNIKL 15
 DB 87 RPLWIFSGNNIKL 101

RESULT 3

S26211
 pectate lyase (EC 4.2.2.2) - common tobacco
 C/Species: Nicotiana tabacum (common tobacco)
 C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 21-Jul-2000
 C/Accession: S26211; S26212; S21933; S22753; S22754
 R/Rogers, H.J.; Harvey, A.; Lonsdale, D.M.
 Plant Mol. Biol. 20, 493-502, 1992

A/Title: Isolation and characterization of a tobacco gene with homology to pectate lyase
 A/Reference number: S26211; MUID:93043039; PMID:1421152

A/Accession: S26211
 A/Molecule type: DNA
 A/Residues: 1-397 <ROG>
 A/Cross-references: EMBL:X67158; NID:G19907; PIDN:CAA47630.1; PID:G19908
 A/Accession: S26212
 A/Molecule type: mRNA
 A/Residues: 119-155; 'C', 157-188, 190, 'G', 191-199, 'D', 200, 'R', 203-248, 'N', 250-381 <RO2>
 A/Cross-references: EMBL:X67159; NID:G19909; PIDN:CAA47631.1; PID:G3980174
 A/Note: translation of the nucleotide sequence is not complete
 R/Lonsdale, D.M.
 submitted to the EMBL Data Library, July 1991

A/Reference number: S21933

A/Accession: S21933
 A/Molecule type: DNA
 A/Residues: 1-397 <LON>
 A/Cross-references: EMBL:X61102; NID:G19981; PIDN:CAA43414.1; PID:G19982
 C/Genetics:

A/Introns: 193/1; 293/2
 A/Suprafamily: pectate lyase LAT59
 C/Keywords: carbon-oxygen lyase

Query Match 71.6%; Score 58; DB 2; Length 397;
 Best Local Similarity 78.6%; Pred. No. 0.029;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RPLWIFSGNNIKL 15
 DB 112 RPLWIFSGNNIKL 125

RESULT 4

T06728
 pectate lyase (EC 4.2.2.2) F28P10.100 - Arabidopsis thaliana
 N/Alternate names: protein F28P10.100
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
 C/Accession: T06728

R/Querier, F.; Choise, N.; Robert, C.; Brotier, P.; Wincker, P.; Catolico, L.; Artigou
 submitted to the Protein Sequence Database, April 1999

A/Reference number: Z15793

A/Accession: T06728

A/Molecule type: DNA

A/Residues: 1-542 <QUB>
 A/Cross-references: EMBL:AL049655
 A/Experimental source: cultivar Columbia; BAC clone F28P10

C/Genetics:

A/Map position: 3
 A/Introns: 46/2; 346/3; 413/2; 480/2
 A/Note: F28P10.100
 C/Suprafamily: pectate lyase LAT59

C/Keywords: carbon-oxygen lyase

Query Match 70.4%; Score 57; DB 2; Length 542;
 Best Local Similarity 71.4%; Pred. No. 0.061;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RPLWIFSGNNIKL 15
 DB 146 RPLWIFSGNNIKL 159

RESULT 5

T00856
 pectate lyase (EC 4.2.2.2) T20P6.14 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001
 C/Accession: T00856; A84440
 R/Rounsley, S.D.; Lin, X.; Ketchum, R.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K
 submitted to the EMBL Data Library, March 1998

A/Description: Arabidopsis thaliana chromosome II BAC T20P6 genomic sequence.
 A/Reference number: Z14206
 A/Accession: T00856
 A/Status: translated from GB/EMBL/DBJ

A/Residues: 1-455 <ROU>
 A/Molecule type: DNA

A/Cross-references: EMBL:AC002521; NID:G2947056; PIDN:AAC05350.1; PID:G2947069

A/Experimental source: cultivar Columbia
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umeyama, L.; Tallon,
 euse, D.; Niekman, W.C.; White, O.; Eism, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: A84440
 A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-455 <STO>

A/Cross-references: GB:AE002093; NID:G2947069; PIDN:AAC05350.1; GSPDB:GND0139

C/Genetics:

A/Map position: 2

A/Introns: 66/2; 295/3; 376/3

A/Note: T20P6.14

C/Suprafamily: pectate lyase LAT59
 C/Keywords: carbon-oxygen lyase

Query Match 69.1%; Score 56; DB 2; Length 455;
 Best Local Similarity 73.3%; Pred. No. 0.075;
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPLWIFSGNNIKL 15
 DB 166 RPLWIFSGNNIKL 180

RESULT 6
 T46165
 pectate lyase-like protein - Arabidopsis thaliana
 N/Alternate names: protein T46165
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
 C/Accession: T46165

R/Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Welcheselgartner, F

submitted to the Protein Sequence Database, December 1999

A/Reference number: Z23025

A/Accession: T46165

A/Status: preliminary
 A/Molecule type: DNA

A/Residues: 1-463 <NTA>
 A/Cross-references: EMBL:AL132958
 A/Experimental source: cultivar Columbia; BAC clone T462
 C/Genetics:
 A/Map position: 3

A:Introns: 27/2; 112/2; 182/1; 265/3; 329/3; 399/2
 A>Note: T4D2.120
 C:Superfamily: pectate lyase LAT59

Query Match 69.1%; Score 56; DB 2; Length 463;
 Best Local Similarity 71.4%; Pred. No. 0.077;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PLWIFSGNNMIXL 15
 DB 120 PLWIFPSNNMIXL 133

RESULT 7

S27098
 Pectate lyase (EC 4.2.2.2) LAT59 - tomato
 N/Alternate names: protein P59
 C/Species: Lycopersicon esculentum (tomato)
 C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 20-Apr-2001
 C/Accession: S27098; S08064
 R/Griffith, I.J.; Yamaguchi, J.; Larrabell, S.K.; Ursin, V.M.; McCormick, S.
 Plant Mol. Biol. 14, 17-28, 1990
 A>Title: Molecular and genetic characterization of two pollen-expressed genes that have
 A/Reference number: S08383; MUID:91322485; PMID:1983191
 A/Accession: S27098
 A/Molecule type: DNA
 A/Residues: 1-448 <WIN>
 A/Cross-references: EMBL:X15499; NID:919270; PIDN:CAA33523.1; PID:919271
 C/Genetics:
 A/Map position: 3
 A/Introns: 67/2; 177/2
 C:Superfamily: pectate lyase LAT59
 C/Keywords: carbon-oxygen lyase

Query Match 67.9%; Score 55; DB 2; Length 449;
 Best Local Similarity 71.4%; Pred. No. 0.11;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PLWIFSGNNMIXL 15
 DB 165 PLWIFKRGNNMIXL 178

RESULT 8

B53240
 allergen Amb a 1.2 precursor - common ragweed
 C/Species: Ambrosia artemisiifolia (common ragweed)
 C/Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 18-Sep-1998
 C/Accession: B53240
 R/Griffith, I.J.; Pollock, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.
 Int. Arch. Allergy Appl. Immunol. 96, 296-304, 1991
 A>Title: Sequence polymorphism of Amb a I and Amb a II, the major allergens in Ambrosia
 A/Reference number: A53240; MUID:92234570; PMID:1809687
 A/Accession: B53240
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-398 <GRI>
 A/Cross-references: GB:M0559
 C:Superfamily: pectate lyase LAT59
 C/Keywords: pollen

Query Match 66.7%; Score 54; DB 2; Length 398;
 Best Local Similarity 73.3%; Pred. No. 0.15;
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPLWIFSGNNMIXL 15
 DB 115 RPLWIFKRGNNMIXL 129

RESULT 9
 B39099
 allergen Amb a 1.2 - common ragweed

C/Species: Ambrosia artemisiifolia (common ragweed)
 C/Date: 27-Nov-1991 #sequence_revision 03-Apr-1992 #text_change 21-Jul-2000
 C/Accession: B39099
 R/Refnar, T.; Griffith, I.J.; Kuo, M.; Bond, J.F.; Rogers, B.L.; Klapper, D.G.
 J. Biol. Chem. 266, 1229-1236, 1991
 A>Title: Cloning of Amb a I (antigen E), the major allergen family of short ragweed
 A/Reference number: A39099; MUID:91093235; PMID:1702434
 A/Accession: B39099
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-398 <RAF>
 A/Cross-references: GB:M62981; NID:9166436; PIDN:AAA32666.1; PID:9166437
 C:Superfamily: pectate lyase LAT59
 C/Keywords: pollen

Query Match 66.7%; Score 54; DB 2; Length 398;
 Best Local Similarity 73.3%; Pred. No. 0.15;
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPLWIFSGNNMIXL 15
 DB 115 RPLWIFKRGNNMIXL 129

RESULT 10

D53240
 allergen Amb a 1.4 precursor - common ragweed
 C/Species: Ambrosia artemisiifolia (common ragweed)
 C/Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 29-Sep-1999
 C/Accession: D53240
 R/Griffith, I.J.; Pollock, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.
 Int. Arch. Allergy Appl. Immunol. 96, 296-304, 1991
 A>Title: Sequence polymorphism of Amb a I and Amb a II, the major allergens in Ambrosia
 A/Reference number: A53240; MUID:92234570; PMID:1809687
 A/Accession: D53240
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-392 <GRI>
 A/Cross-references: GB:M0562; NID:9166444; PIDN:AAA32670.1; PID:9166445
 C:Superfamily: pectate lyase LAT59
 C/Keywords: pollen

Query Match 65.4%; Score 53; DB 2; Length 392;
 Best Local Similarity 66.7%; Pred. No. 0.21;
 Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPLWIFSGNNMIXL 15
 DB 114 RPLWIFKRGNNMIXL 128

RESULT 11

T51456
 Pectate lyase-like protein - Arabidopsis thaliana
 N/Alternate names: protein F2G14_230
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 17-Nov-2000
 C/Accession: T51456
 R/Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; et al.
 submitted to the Protein Sequence Database, August 2000
 A/Reference number: Z25394
 A/Accession: T51456
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-472 <SAT>
 A/Cross-references: EMBL:AL391146
 A/Experimental source: cultivar Columbia, BAC clone F2G14
 C/Genetics:
 A/Map position: 5
 A/Introns: 65/2; 316/3; 398/3
 A/Note: F2G14_230
 C:Superfamily: pectate lyase LAT59

Query Match 65.4%; Score 53; DB 2; Length 472;
 Best Local Similarity 64.3%; Pred. No. 0.26;
 Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 RPLMIIFSGNNIKL 15
 ||||| :|||
 Db 188 RPLMIIFKNDWVIRL 201

RESULT 12

A:Accession: A39099
 A:Species: Ambrosia artemisiifolia (common ragweed)
 C:Date: 27-Nov-1991 #sequence_revision 03-Apr-1992 #text_change 29-Sep-1999
 C:Accession: A39099; A60895; A53240
 R:Reimer, T.; Griffith, I.J.; Kuo, M.; Bond, J.F.; Rogers, B.L.; Klapper, D.G.
 J. Biol. Chem. 266, 1229-1236, 1991
 A:Title: Cloning of Amb a I (antigen E), the major allergen family of short ragweed pollen
 A:Reference number: A39099; PMID:91093235; PMID:1702434
 A:Accession: A39099
 A:Molecule type: mRNA
 A:Residues: 1-397 <RA>
 A:Cross-references: GB:M63116
 R:Smith, J.J.; Olson, J.R.; Klapper, D.G.
 Mol. Immunol. 25, 355-365, 1988
 A:Title: Monoclonal antibodies to denatured ragweed pollen allergen Amb a I: characterization
 A:Reference number: A60895; PMID:88288254; PMID:2456454
 A:Accession: A60895
 A:Molecule type: protein
 A:Residues: 256-273; 292-303; 'W', 305-306 <SMI>
 R:Griffith, I.J.; Pollock, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.
 Int. Arch. Allergy Appl. Immunol. 96, 296-304, 1991
 A:Title: Sequence polymorphism of Amb a I and Amb a II, the major allergens in Ambrosia
 A:Reference number: A53240; PMID:92234570; PMID:1809687
 A:Accession: A53240
 A:Molecule type: mRNA
 A:Status: preliminary
 A:Residues: 1-396 <GRI>
 A:Cross-references: GB:M80558; NID:G166434; PIDN:AAA32665.1; PID:G166435
 C:Superfamily: pectate lyase LATS9
 C:Keywords: glycoprotein; pollen
 F:36/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 64.2%; Score 52; DB 2; Length 396;
 Best Local Similarity 66.7%; Pred. No. 0.32;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPLMIIFSGNNIKL 15
 ||||| :|||
 Db 113 RPLMIIFKNDWVIRL 127

RESULT 13

C53240
 allergen Amb a I.3 precursor - common ragweed
 C:Species: Ambrosia artemisiifolia (common ragweed)
 C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 18-Sep-1998
 C:Accession: C53240
 R:Griffith, I.J.; Pollock, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.
 Int. Arch. Allergy Appl. Immunol. 96, 296-304, 1991
 A:Title: Sequence polymorphism of Amb a I and Amb a II, the major allergens in Ambrosia
 A:Reference number: A53240; PMID:92234570; PMID:1809687
 A:Accession: C53240
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-397 <GRI>
 A:Cross-references: GB:M80560
 C:Superfamily: pectate lyase LATS9
 C:Keywords: pollen

Query Match 64.2%; Score 52; DB 2; Length 397;
 Best Local Similarity 66.7%; Pred. No. 0.32;

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RPLMIIFSGNNIKL 15
 ||||| :|||
 Db 114 RPLMIIFKNDWVIRL 128

RESULT 14

C39099
 allergen Amb a I.3 - common ragweed
 C:Species: Ambrosia artemisiifolia (common ragweed)
 C:Date: 27-Nov-1991 #sequence_revision 03-Apr-1992 #text_change 29-Sep-1999
 C:Accession: C39099
 R:Reimer, T.; Griffith, I.J.; Kuo, M.; Bond, J.F.; Rogers, B.L.; Klapper, D.G.
 J. Biol. Chem. 266, 1229-1236, 1991
 A:Title: Cloning of Amb a I (antigen E), the major allergen family of short ragweed po
 A:Reference number: A39099; PMID:91093235; PMID:1702434
 A:Accession: C39099
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-397 <RA>
 A:Cross-references: GB:M62961; NID:G166440; PIDN:AAA32668.1; PID:G166441
 C:Superfamily: pectate lyase LATS9
 C:Keywords: pollen

Query Match 64.2%; Score 52; DB 2; Length 397;
 Best Local Similarity 66.7%; Pred. No. 0.32;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPLMIIFSGNNIKL 15
 ||||| :|||
 Db 114 RPLMIIFKNDWVIRL 128

RESULT 15

S29612
 pectate lyase (EC 4.2.2.2) - trumpet lily
 C:Species: Lilium longiflorum (trumpet lily)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Sep-1999
 C:Accession: S29612
 R:Kim, S.; Finkel, D.J.; An, G.
 submitted to the EMBL Data Library, October 1992
 A:Description: Abundance patterns of lily pollen cDNAs: characterization of three pol
 A:Reference number: S29611
 A:Accession: S29612
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-434 <KIM>
 A:Cross-references: EMBL:Z17328; NID:G19450; PIDN:CAA78976.1; PID:G19451
 A:Experimental source: cv. Nellie White, mature flower
 C:Superfamily: pectate lyase LATS9
 C:Keywords: carbon-oxygen lyase

Query Match 64.2%; Score 52; DB 2; Length 434;
 Best Local Similarity 66.7%; Pred. No. 0.35;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPLMIIFSGNNIKL 15
 ||||| :|||
 Db 151 RPLMIIFKNDWVIRL 165

Search completed: April 20, 2003, 13:15:38
 Job time : 7.07895 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds
(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524d-28

Perfect score: 81

Sequence: 1 RPLWIFSGNMWIKL 15

Scoring table: BLOSUM62

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	81	100.0	374	1 SBP_CRYJA
2	62	76.5	346	1 MPAT_CUPAR
3	62	76.5	367	1 MPAT_JUNAS
4	62	76.5	375	1 MPAT_CHAAB
5	58	71.6	397	1 PEL_TOBAC
6	55	67.9	449	1 PERS_LYCES
7	54	66.7	398	1 MP12_AMBAR
8	53	65.4	392	1 MP14_AMBAR
9	52	64.2	396	1 MP11_AMBAR
10	52	64.2	397	1 MP13_AMBAR
11	52	64.2	434	1 PEL_LILLO
12	48	59.3	397	1 PEA2_AMBAR
13	47	58.0	385	1 YEHY_ECOLI
14	46	56.8	398	1 PERS_LYCES
15	46	56.8	404	1 PERS_LYCES
16	41	50.6	915	1 ATCU_VIBCH
17	40	49.4	202	1 KGRU_SCHPO
18	40	49.4	753	1 PERS_HUMAN
19	39.5	48.8	1043	1 P11D_MOUSE
20	39	48.1	333	1 YH98_ARCFU
21	39	48.1	372	1 Y412_MYCPN
22	39	48.1	440	1 OMGP_HUMAN
23	39	48.1	440	1 OMGP_MOUSE
24	39	48.1	2504	1 FAS_HUMAN
25	39	48.1	2505	1 FAS_MOUSE
26	38.5	47.5	1887	1 RPB1_DROME
27	38	46.9	202	1 T2MU_MYCSP
28	38	46.9	227	1 SRPB_SCHPO
29	38	46.9	254	1 PHSB_SALT
30	38	46.9	475	1 AMT2_ARATH
31	38	46.9	538	1 RKP1_RHIME
32	38	46.9	2511	1 FAS_CHICK
33	37.5	46.3	1852	1 RPB1_CABEL

34	37	45.7	180	1	YPSA_BACSU	P50838 bacillus su
35	37	45.7	296	1	YFMW_BACSU	P39649 bacillus su
36	37	45.7	351	1	Y630_RICCR	O92ct3 rickettsia
37	37	45.7	386	1	KPBQ_HUMAN	Q16816 homo sapien
38	37	45.7	387	1	KPBQ_MOUSE	P07934 mus musculu
39	37	45.7	387	1	KPBQ_MOUSE	P13286 ratu mus norv
40	37	45.7	457	1	AROP_ECOLI	P15993 escherichia
41	37	45.7	588	1	CAR7_CANAL	P43096 candida alb
42	37	45.7	656	1	VEXE_SALT	P43112 salmonella
43	37	45.7	851	1	BP2_CABEL	P29691 caenorhabd
44	37	45.7	862	1	VG01_HSV1	Q00132 ictaluri h
45	36.5	45.1	267	1	CYSH_PSEAE	O05927 pseudomonas

ALIGNMENTS

RESULT 1
ID SBP_CRYJA STANDARD; PRT; 374 AA.
AC P18632;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Sugi basic protein precursor (SBP) (Major allergen Cry j 1) (Cry j I).
OS Cryptomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OK Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OK NCBI_TaxID=3369;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=94183234; PubMed=8135802;
RA Sone T., Komiya N., Shimizu K., Kusabe T., Morikubo K.,
RA Kino K.;
RT "Cloning and sequencing of cDNA coding for Cry j I, a major allergen
RT of Japanese cedar pollen.";
RN Biochem. Biophys. Res. Commun. 199;619-625(1994).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=94183234; PubMed=8135802;
RA Sone T., Komiya N., Shimizu K., Kusabe T., Morikubo K.,
RA Kino K.;
RT "Cloning and sequencing of cDNA coding for Cry j I, a major allergen
RT of Japanese cedar pollen.";
RN Biochem. Biophys. Res. Commun. 199;619-625(1994).
[3]
RP SEQUENCE OF 22-41.
RC TISSUE=Pollen;
RX MEDLINE=89031257; PubMed=3181436;
RA Tanai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inoue S.,
RA Matsumoto T.;
RT "N-terminal amino acid sequence of a major allergen of Japanese cedar
RT pollen (Cry j I).";
RN FEBS Lett. 239;332-332(1988).
[4]
RP CARBOHYDRATES.
RC TISSUE=Pollen;
RX MEDLINE=95003748; PubMed=7920021;
RA Hijikata A., Matsumoto I., Kojima K., Ogawa H.;
RT "Antigenicity of the oligosaccharide moiety of the Japanese cedar
RT (Cryptomeria japonica) pollen allergen, Cry j I.";
RN Int. Arch. Allergy Immunol. 105;198-202(1994).
[5]
RP STRUCTURE OF CARBOHYDRATES.
RC TISSUE=Pollen;
RX MEDLINE=9532249; PubMed=7608114;
RA Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Usui M.,
RA Fukuda S., Hanzawa H., Hanyama H., Kurimoto M.;
RT "Carbohydrate structures of the glycoprotein allergen Cry j I from
RT Japanese cedar (Cryptomeria japonica) pollen.";
RL J. Biochem. 117;289-295(1995).
CC -!- PTM: CONTAINS FUCOSE/Xylose-CONTAINING N-LINKED OLIGOSACCHARIDES.
CC -!- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
CC POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN.
CC -!- MISCELLANEOUS: THE SEQUENCE OF CRY J I FORM A IS SHOWN HERE. FORM

FT SIGNAL 1 21
 FT CHAIN 22 367
 FT CARBOHYD 148 148
 FT CARBOHYD 178 178
 SQ SEQUENCE 367 AA; 39824 MW; FC9B81E675662E49 CRC64;

Query Match
 Best Local Similarity 76.5%; Score 62; DB 1; Length 367;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 RPLWIFSGNNIKL 15
 DB 87 KALWIFSGNNIKL 101

RESULT 4
 ID MPAL CHAOB STANDARD; PRT; 375 AA.
 AC 096385;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major pollen allergen Cha o 1 precursor.
 OS Chamaecyparis obtusa (Japanese cypress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
 CC Chamaecyparis.

OK NCBI_TaxID=13415;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Pollen;
 RX MEDLINE=96265194; PubMed=8676896;

RA Suzuki M., Komiyama N., Itoh M., Sone T., Kuno K., Takagi I.,
 Ohta N.;
 RT "Purification, characterization and molecular cloning of Cha o 1, a
 major allergen of Chamaecyparis obtusa (Japanese cypress) pollen.";
 RL Mol. Immunol. 33:451-460(1996)
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; D45404; BAA08246.1; -
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KW Allergen; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 375
 FT CARBOHYD 110 110
 FT CARBOHYD 148 148
 FT CARBOHYD 178 178
 FT CARBOHYD 293 293
 FT CARBOHYD 352 352
 SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF7066DBF CRC64;

Query Match
 Best Local Similarity 76.5%; Score 62; DB 1; Length 375;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 RPLWIFSGNNIKL 15
 DB 87 RSLWIFSGNNIKL 101

RESULT 5
 PEL_TOBAC

ID PEL_TOBAC STANDARD; PRT; 397 AA.
 AC P40572;

DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Pectate lyase precursor (EC 4.2.2.2).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asceridae; easterids I; Solanales; Solanaceae; Nicotiana.
 OK NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=cv. Samsun; Tissue=Pollen;
 RX MEDLINE=93043039; PubMed=1421152;
 RA Rogers H.J., Harvey A., Lonsdale D.M.;
 RT "Isolation and characterization of a tobacco gene with homology to
 RT pectate lyase which is specifically expressed during
 RT microsporogenesis.";
 RL Plant Mol. Biol. 20:493-502(1992).

CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC -1- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED LATE IN POLLEN
 CC DEVELOPMENT.

CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X67158; CAA47630.1; -
 DR EMBL; X67159; CAA47631.1; -
 DR EMBL; X61102; CAA43414.1; -
 DR PIR; S26211; S26211.
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.

KW Lyase; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 397
 FT ACT SITE 272
 FT CARBOHYD 134 134
 FT CARBOHYD 227 227
 FT CARBOHYD 156 156
 FT CONFLICT 189 190
 FT CONFLICT 200 200
 FT CONFLICT 202 202
 FT CONFLICT 249 249
 SQ SEQUENCE 397 AA; 44351 MW; EF0A82CB5DA7643F CRC64;

Query Match
 Best Local Similarity 71.6%; Score 58; DB 1; Length 397;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PLWIFSGNNIKL 15
 DB 112 PLWIFSGNNIKL 125

RESULT 6
 ID PEL59_LYCES STANDARD; PRT; 449 AA.
 AC P15722;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Probable pectate lyase P59 precursor (EC 4.2.2.2).
 GN LATS9.

OS Lycopersicon esculentum (Tomato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asterales; easterids I; Solanales; Solanaceae; Solanum.
 CC NCBI_TaxID=4081;
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV_VF36; TISSUE=Anther;
 RX MEDLINE=9132485; PubMed=1983191;
 RA Wing R.A., Yamaguchi J., Larebell S.K., Ursin V.M., McCormick S.;
 RT "Molecular and genetic characterization of two pollen-expressed genes
 RT that have sequence similarity to pectate lyases of the plant pathogen
 RT Erwinia";
 RT Plant Mol. Biol. 14:17-28(1990).
 CC
 CC -1- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE
 CC GROWTH.
 CC
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
 CC
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X15499; CA33523.1; -;
 DR PIR; S27098; S27098.
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 DR Lyase; Multigene family; Signal.
 KM
 FT SIGNAL 22
 FT CHAIN 23 449 PROBABLE PECTATE LYASE P59.
 FT ACT SITE 325 325 POTENTIAL.
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 449 AA; 50893 MW; 17E3AA13F173B03C CRC64;
 QY Query Match 67.9%; Score 55; DB 1; Length 449;
 Best Local Similarity 71.4%; Pred. No. 0.046;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 DB 165 PLWIIIFKGMNIRL 178
 RESULT 7
 MP12_AMBAR STANDARD; PRT; 398 AA.
 ID MP12_AMBAR
 AC P27760;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pollen allergen Amb a 1.2 precursor (Antigen B) (Antigen Amb a I).
 OS Ambrosia artemisiifolia (Short ragweed).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asterales; easterids II; Asterales; Asteraceae; Asteroideae;
 CC Helianthaceae; Ambrosia.
 CC NCBI_TaxID=4212;
 RN
 RN SEQUENCE FROM N.A.
 RC TISSUE=Pollen;
 RX MEDLINE=91093235; PubMed=1702434;
 RA Ratnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
 RA Klapper D.G.;

RT "Cloning of Amb a I (antigen E), the major allergen family of short
 RT ragweed pollen.";
 RT J. Biol. Chem. 266:1229-1236(1991).
 RN
 RN SEQUENCE FROM N.A., AND VARIANTS.
 RP TISSUE=Pollen;
 RX MEDLINE=92234570; PubMed=1809687;
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
 RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
 RT in Ambrosia artemisiifolia (short ragweed).";
 RT Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
 CC
 CC -1- SUBUNIT: MONOMER.
 CC
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
 CC
 CC -1- PTM: The N-terminus is blocked.
 CC
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
 CC
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M62981; AAA32666.1; -;
 DR EMBL; M80559; AAA32667.1; -;
 DR PIR; B39099; B39099.
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 DR Antigen; Allergen; Signal; Multigene family; Polymorphism.
 KM
 FT SIGNAL 1 25
 FT CHAIN 26 398 POTENTIAL.
 FT VARIANT 345 345 POLLEN ALLERGEN AMB A 1.2.
 FT VARIANT 381 381 R -> K.
 SQ SEQUENCE 398 AA; 43664 MW; 020DC662D9B7416C CRC64;
 QY Query Match 66.7%; Score 54; DB 1; Length 398;
 Best Local Similarity 73.3%; Pred. No. 0.06;
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 DB 115 RPLWIIIFKGMNIRL 129
 RESULT 8
 MP14_AMBAR STANDARD; PRT; 392 AA.
 ID MP14_AMBAR
 AC P28744;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pollen allergen Amb a 1.4 precursor (Antigen B) (Antigen Amb a I).
 OS Ambrosia artemisiifolia (Short ragweed).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asterales; easterids II; Asterales; Asteraceae; Asteroideae;
 CC Helianthaceae; Ambrosia.
 CC NCBI_TaxID=4212;
 RN
 RN SEQUENCE FROM N.A., AND VARIANTS.
 RC TISSUE=Pollen;
 RX MEDLINE=92234570; PubMed=1809687;
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
 RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
 RT in Ambrosia artemisiifolia (short ragweed).";
 RT Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
 CC
 CC -1- SUBUNIT: MONOMER.
 CC
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
 CC
 CC -1- PTM: The N-terminus is blocked.

CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M80562; AAA32670.1; --
 CC PIR; D53240; D53240.
 CC InterPro; IPR002022; Amb allergen.
 CC Pfam; PF00544; pec_lyase; 1.
 CC PRINTS; PRO0807; AMBALLERGEN.
 CC KX Antigen; Allergen; Signal; Multigene family; Polymorphism.
 CC FT SIGNAL 1 25 POTENTIAL.
 CC CHAIN 26 392 POLLEN ALLERGEN AMB A 1.4.
 CC VARIANT 182 188 SHDGPV -> CNDGPPA.
 CC SEQUENCE 392 AA; 42842 MW; 7B6219C12F365DA9 CRC64;
 CC -----
 CC Query Match 65.4%; Score 53; DB 1; Length 392;
 CC Best Local Similarity 66.7%; Pred. No. 0.088;
 CC Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 CC -----
 CC QY 1 RPLWIFSGNNIKL 15
 CC DB 114 RPLWIFARDVIRL 128
 CC -----
 CC RESULT 9
 CC ID MP13_AMBAR STANDARD; PRT; 396 AA.
 CC AC P27759;
 CC DT 01-AUG-1992 (Rel. 23, Created)
 CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Pollen allergen Amb a 1.1 precursor (Antigen E) (Antigen Amb a I).
 CC OS Ambrosia artemisiifolia (short ragweed).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 CC OC Helianthaceae; Ambrosia.
 CC OX NCBI_TaxID=4212;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC RC TISSUE=Pollen;
 CC RX MEDLINE=91093235; PubMed=1702434;
 CC RA Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
 CC RA Klapper D.G.;
 CC RT "Cloning of Amb a I (antigen E), the major allergen family of short
 CC RT ragweed pollen.";
 CC RL J. Biol. Chem. 266:1229-1236(1991).
 CC RN [2]
 CC RP SEQUENCE FROM N.A., AND VARIANTS.
 CC RC TISSUE=Pollen;
 CC RX MEDLINE=92234570; PubMed=1809687;
 CC RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
 CC RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
 CC RT in Ambrosia artemisiifolia (short ragweed).";
 CC RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
 CC RN [1] SUBUNIT: MONOMER.
 CC CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
 CC CC -1- PTM: The N-terminus is blocked.
 CC CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
 CC CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M63116; -; NOT ANNOTATED_CDS.
 CC DR EMBL; M80558; AAA32655.1; --
 CC DR PIR; A39099; A39099.
 CC DR PIR; A53240; A53240.
 CC InterPro; IPR002022; Amb allergen.
 CC Pfam; PF00544; pec_lyase; 1.
 CC PRINTS; PRO0807; AMBALLERGEN.
 CC KX Antigen; Allergen; Signal; Multigene family; Polymorphism.
 CC FT SIGNAL 1 25 POTENTIAL.
 CC CHAIN 26 396 POLLEN ALLERGEN AMB A 1.1.
 CC VARIANT 92 92 E -> D.
 CC SEQUENCE 396 AA; 42709 MW; 0CE7DCEB3B8841D CRC64;
 CC -----
 CC Query Match 64.2%; Score 53; DB 1; Length 396;
 CC Best Local Similarity 66.7%; Pred. No. 0.13;
 CC Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 CC -----
 CC QY 1 RPLWIFSGNNIKL 15
 CC DB 113 RPLWIFERDVIRL 127
 CC -----
 CC RESULT 10
 CC ID MP13_AMBAR STANDARD; PRT; 397 AA.
 CC AC P27761;
 CC DT 01-AUG-1992 (Rel. 23, Created)
 CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Pollen allergen Amb a 1.3 precursor (Antigen E) (Antigen Amb a I).
 CC OS Ambrosia artemisiifolia (short ragweed).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 CC OC Helianthaceae; Ambrosia.
 CC OX NCBI_TaxID=4212;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Pollen;
 CC RX MEDLINE=91093235; PubMed=1702434;
 CC RA Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
 CC RA Klapper D.G.;
 CC RT "Cloning of Amb a I (antigen E), the major allergen family of short
 CC RT ragweed pollen.";
 CC RL J. Biol. Chem. 266:1229-1236(1991).
 CC RN [2]
 CC RP SEQUENCE FROM N.A., AND VARIANTS.
 CC RC TISSUE=Pollen;
 CC RX MEDLINE=92234570; PubMed=1809687;
 CC RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
 CC RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
 CC RT in Ambrosia artemisiifolia (short ragweed).";
 CC RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
 CC RN [1] SUBUNIT: MONOMER.
 CC CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
 CC CC -1- PTM: The N-terminus is blocked.
 CC CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
 CC CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----


```

DR EMBL; M62961; AAA32668.1; -.
DR EMBL; M80560; AAA32669.1; ALT_SEQ.
DR PIR; C39099; C39099.
DR InterPro; IPR002022; Amb allergen.
DR Pfam; PF00544; pec_lyase_1.
DR PRINTS; PR00807; AMBALLERGEN.
KW Antigen; Allergen; Signal; Multigene family; Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 397 POTENTIAL.
FT VARIANT 48 48 L -> Y.
SQ SEQUENCE 397 AA; 42928 MW; C8DB41257590DDA CRC64;

Query Match 64.2%; Score 52; DB 1; Length 397;
Best Local Similarity 66.7%; Pred. No. 0.13;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPLMTIFSGNNIKL 15
DB 114 RPLMTIFKNDVINTL 128

RESULT 11
ID PEL_LILLO STANDARD; PRT; 434 AA.
AC P40973;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Pectate lyase precursor (EC 4.2.2.2).
OS Liliaceae; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
OC Liliaceae;
OC NCBI_TaxID=4690;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nellie white; TISSUE=pollen;
RA Kim S.R., Finkel D.J., An G.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Elimination of pectate to give
CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
CC their non-reducing ends.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL; Z17328; CAA78976.1; -.
CC DR EMBL; L18911; AAA33398.1; -.
CC DR PIR; S29612; S29612.
CC DR InterPro; IPR002022; Amb allergen.
CC Pfam; PF00544; pec_lyase_1.
CC PRINTS; PR00807; AMBALLERGEN.
KW Antigen; Signal; Polymorphism.
FT SIGNAL 1 22
FT CHAIN 23 434 POTENTIAL.
FT ACT SITE 312 312 POTENTIAL.
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 434 AA; 48457 MW; C1F3E30AD2BBD064 CRC64;

Query Match 64.2%; Score 52; DB 1; Length 434;
Best Local Similarity 66.7%; Pred. No. 0.15;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPLMTIFSGNNIKL 15
DB 151 RPLMTIFKNDVINTL 165

```

```

RESULT 12
ID MP22_AMBAR STANDARD; PRT; 397 AA.
AC P2762;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pollen allergen Amb a 2 precursor (antigen K) (Antigen Amb a II).
OS Ambrosia artemisiifolia (short ragweed); Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; eudicotyledons; core eudicots;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; easterids II; Asterales; Asteraceae; Asteroideae;
OC Helianthaceae; Ambrosia.
OC NCBI_TaxID=4212;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=flower;
RC MEDLINE=92013060; PubMed=1717566;
RA Rogers B.L., Morgenstern J.P., Griffith I.J., Yu X.-B.,
RA Connors C.M., Brauer A.W., King T.P., Garman R.D., Kuo M.-C.C.;
RT "Complete sequence of the allergen Amb alpha II. Recombinant
RT expression and reactivity with T cells from ragweed allergic
RT patients."
RL J. Immunol. 147:2547-2552 (1991).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSUE=pollen;
RC MEDLINE=92234570; PubMed=1809687;
RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
RT in Ambrosia artemisiifolia (short ragweed)."
RL Int. Arch. Allergy Appl. Immunol. 96:296-304 (1991).
CC -1- SUBUNIT: MONOMER.
CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC -1- PTM: The N-terminus is blocked.
CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY J I SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL; M80561; AAA32671.1; -.
CC DR PIR; A46469; A46469.
CC DR InterPro; IPR002022; Amb allergen.
CC Pfam; PF00544; pec_lyase_1.
CC PRINTS; PR00807; AMBALLERGEN.
KW Antigen; Allergen; Signal; Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 397
FT VARIANT 70 70 N -> D (DETECTED ONLY IN FLOWER DNA).
FT VARIANT 138 138 K -> T (DETECTED ONLY IN FLOWER DNA).
FT VARIANT 321 321 K -> R (DETECTED ONLY IN FLOWER DNA).
SQ SEQUENCE 397 AA; 44082 MW; C7861784C9A3DDDD CRC64;

Query Match 59.3%; Score 48; DB 1; Length 397;
Best Local Similarity 66.7%; Pred. No. 0.66;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPLMTIFSGNNIKL 15
DB 114 RPLMTIFKNDVINTL 128

RESULT 13
YEHY_ECOLI

```


ID YEHY_ECOLI STANDARD; PRT; 385 AA.
 AC P3361; P76435; Rel. 28, Created
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Hypothetical ABC transporter permease protein yehy.
 OS YEHY OR B2130.
 OC Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_Taxid=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / BHD2600;
 RA Richerich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
 RL Church G.M.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MGI655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.U.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM. PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE
 CC SUBSTRATE ACROSS THE MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -1- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM PERMEASE FAMILY. CYSTW SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U00007; AAA60493.1; ALT_INT.
 DR EMBL; AE000302; AAC75191.1; -.
 DR Ecogene; EG12011; Yehy.
 DR InterPro; IPR00515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 DR PROSITE; PS00402; BPD_TRANSP_INN_MEMBER; 1.
 KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
 KM Complete proteome.
 FT TRANSMEM 6 26 POTENTIAL.
 FT TRANSMEM 45 65 POTENTIAL.
 FT TRANSMEM 70 90 POTENTIAL.
 FT TRANSMEM 107 127 POTENTIAL.
 FT TRANSMEM 181 201 POTENTIAL.
 FT TRANSMEM 225 245 POTENTIAL.
 FT TRANSMEM 248 268 POTENTIAL.
 FT TRANSMEM 323 343 POTENTIAL.
 FT TRANSMEM 347 367 POTENTIAL.
 FT TRANSMEM 346 385 POTENTIAL.
 FT CONFLICT 346 385
 LLSAIDVILGVIPVITAVTDLFDLLMLKVRND
 -> AKORH (IN REF. 1).
 SQ SEQUENCE 385 AA; 41138 MW; E57055E3A2B141CC CRC64;
 Query Match 58.0%; Score 47; DB 1; Length 385;
 Best Local Similarity 54.5%; Pred. No. 0.95;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

RESULT 14
 ID P56_LYCES STANDARD; PRT; 398 AA.
 AC P56_LYCES
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 GN Probable pectate lyase P56 precursor (EC 4.2.2.2).
 OS Latis.
 OC Lycopodium esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_Taxid=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. VF36; TISSUE=Anther;
 RX MEDLINE=91322485; PubMed=1983191;
 RA Wing R.A., Yamaguchi J., Larabell S.K., Ursin V.M., McCormick S.,
 RT "Molecular and genetic characterization of two pollen-expressed genes
 RT that have sequence similarity to pectate lyases of the plant pathogen
 RT Erwinia.";
 RL Plant Mol. Biol. 14:17-28(1990).
 [2]
 RP REVISIONS.
 RC STRAIN=cv. VF36; TISSUE=Anther;
 RA Wing R.A.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE
 CC GROWTH.
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X15500; CA33524.1; -.
 DR PIR; S08383; S08383.
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KW Lyase; Multigene family; Signal.
 FT SIGNAL 1 27 OR 22 (POTENTIAL).
 FT CHAIN 28 398 PROBABLE PECTATE LYASE P56.
 FT ACT_SITE 273 273 POTENTIAL.
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 398 AA; 44563 MW; 8D676250BD98C7C8 CRC64;
 Query Match 56.8%; Score 46; DB 1; Length 398;
 Best Local Similarity 57.1%; Pred. No. 1.5;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 113 PLWITFARSMRRL 126
 QY 2 PLWITFSGNNIKL 15
 ID 9612_LYCES STANDARD; PRT; 404 AA.
 AC P24396;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

DE   Style development-specific protein 9612 precursor.
GN   9612.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_Taxid=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. VE36; TISSUE=Pistil;
RX   MEDLINE=91117185; PubMed=2277637;
RA   Budellier K.A., Smith A.G., Gasser C.S.;
RT   "Regulation of a stylar transmitting tissue-specific gene in
RT   wild-type and transgenic tomato and tobacco.";
RL   Mol. Gen. Genet. 224:183-192(1990).
CC   -1- FUNCTION: MAY HAVE A ROLE IN THE DEVELOPMENT OF THE TRANSMITTING
CC   -1- TISSUE OF THE STYLE AND/OR IN THE EVENTS RELATED TO POLLINATION
CC   SUCH AS SOME ASPECT IN THE FACILITATION OF COMPATIBLE POLLEN TUBE
CC   GROWTH.
CC   -1- SUBCELLULAR LOCATION: Secreted.
CC   -1- TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN THE PISTIL WHERE IT IS
CC   FOUND IN THE OUTER FIVE LAYERS OF THE STRANDS OF TRANSMITTING
CC   TISSUE WITHIN THE UPPER TWO-THIRDS OF THE STYLE. FOUND AT MUCH
CC   LOWER LEVELS IN THE ANTHERS AND VEGETATIVE ORGANS.
CC   -1- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND DURING ANTHESIS.
CC   -1- SIMILARITY: 54% IDENTICAL TO TOMATO PROTEINS P59 (AC P15722)
CC   AND P56 (AC P15721).
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X55193; CA38979.1; -.
DR   PIR; S12209; S12209.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase_1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Signal.
FT   SIGNAL. 1 20
FT   CHAIN 21 404
FT   CARBOHYD 37 37
FT   CARBOHYD 191 191
SQ   SEQUENCE 404 AA; 44298 MW; B26ED69B128D8675 CRC64;
      N-LINKED (GLCNAC. . .) (POTENTIAL).
      N-LINKED (GLCNAC. . .) (POTENTIAL).
Query Match 56.8%; Score 46; DB 1; Length 404;
Best Local Similarity 64.3%; Pred. No. 1.5;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 PLWIFSGNNIKL 15
   ||||| :|||
Db 120 PLWIFKRDWVIQL 133

```

Search completed: April 20, 2003, 13:07:29
 Job time : 3.92105 secs

GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 12.6711 Seconds

(without alignments)
243,919 Million cell updates/sec

Title: US-09-142-524d-28

Perfect score: 81

Sequence: 1 RPLMTIFSGNNMIXL 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	374	10	Q8RUR1
2	62	76.5	367	10	Q91LT2
3	62	76.5	367	10	Q91LT1
4	62	76.5	367	10	Q9M4S6
5	62	76.5	367	10	Q9M4S5
6	62	76.5	367	10	Q9M4S4
7	62	76.5	367	10	Q9M4S3
8	62	76.5	367	10	Q9M4S2
9	62	76.5	367	10	Q9M4S1
10	62	76.5	367	10	Q9M4S0
11	58	71.6	453	10	Q94LR5
12	57	70.4	501	10	Q93Z04
13	57	70.4	542	10	Q9SV40
14	56	69.1	455	10	Q64510
15	56	69.1	463	10	Q9SCP2
16	53	65.4	472	10	Q9LFP5

17	51	63.0	181	10	Q23666	Q23666 arabidopsis
18	51	63.0	392	10	Q9FXD8	Q9FXD8 arabidopsis
19	51	63.0	408	10	Q8W116	Q8W116 arabidopsis
20	51	63.0	438	10	Q43862	Q43862 zea mays (m
21	51	63.0	459	10	Q23665	Q23665 arabidopsis
22	51	63.0	459	10	Q9M9S2	Q9M9S2 arabidopsis
23	50	61.7	274	10	Q24159	Q24159 nicotiana t
24	50	61.7	354	10	Q940Q1	Q940Q1 arabidopsis
25	50	61.7	358	10	Q9M505	Q9M505 vicia vinit
26	50	61.7	398	10	Q93XJ1	Q93XJ1 salix glig
27	50	61.7	409	10	Q23017	Q23017 arabidopsis
28	50	61.7	452	10	Q9LRM5	Q9LRM5 arabidopsis
29	50	61.7	475	10	Q9SRH4	Q9SRH4 arabidopsis
30	49	60.5	226	10	Q23667	Q23667 arabidopsis
31	49	60.5	404	10	Q9SB71	Q9SB71 arabidopsis
32	49	60.5	408	10	Q9CSM8	Q9CSM8 arabidopsis
33	49	60.5	410	10	Q9FMK5	Q9FMK5 arabidopsis
34	49	60.5	432	10	Q93Z25	Q93Z25 arabidopsis
35	49	60.5	501	16	Q9ZC02	Q9ZC02 streptomyce
36	48	59.3	405	10	Q24416	Q24416 fragaria an
37	48	59.3	405	10	Q94PT6	Q94PT6 fragaria an
38	48	59.3	449	10	Q95SK1	Q95SK1 drosophila
39	48	59.3	622	5	Q95SK1	Q95SK1 drosophila
40	47	58.0	324	10	Q8S345	Q8S345 capsicum an
41	47	58.0	385	16	Q8X670	Q8X670 escherichia
42	47	58.0	392	10	Q9FM66	Q9FM66 arabidopsis
43	47	58.0	409	10	Q9LFT0	Q9LFT0 arabidopsis
44	46	56.8	401	10	Q24554	Q24554 zinnia eleg
45	46	56.8	407	10	Q9SDW4	Q9SDW4 musa acumin

ALIGNMENTS

```

RESULT 1
ID Q8RUR1 PRELIMINARY; PRT; 374 AA.
AC Q8RUR1;
DT 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Cry j 1 precursor.
GN Cry j 1.1 OR Cry j 1.2.
OS Cryptomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta.
OC Spermatophyta; Coniferales; Cupressaceae; Cryptomeria.
OX NCBI_Taxid=3369;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=POLLEN;
RA Futamura N., Shinohara K.;
RT "Isolation and characterization of cDNAs encoding major allergen Cry j
RT 1 from Cryptomeria japonica pollen."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB081309; BAB86286.1; --
DR EMBL; AB081310; BAB86287.1; --
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;
Query Match 100.0%; Score 81; DB 10; Length 374;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cry 1 RPLMTIFSGNNMIXL 15
Db 87 RPLMTIFSGNNMIXL 101
RESULT 2
ID Q91LT2 PRELIMINARY; PRT; 367 AA.

```

AC Q9L12; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pollen major allergen 1-2.
 OS Juniperus virginiana (Eastern red cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 NCBI_TaxID=39584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21315424; PubMed=11422137;
 RA Miodoro-Horvut T.M., Goldblum R.M., Brooks E.G.;
 RT "Identification of mutations in the genes for the pollen allergens of
 eastern red cedar (Juniperus virginiana).";
 RL Clin. Exp. Allergy 31:771-778(2001).
 DR EMBL; AF151427; AAF80164.1; -;
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF89586 CRC64;
 Query Match 76.5%; Score 62; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RPLMIIFSGNNIKL 15
 ID : ||||| |||||
 Db 87 KTLMIIFSGNNIKL 101
 RESULT 3
 Q9L11 PRELIMINARY; PRT; 367 AA.
 ID Q9L11;
 AC Q9L11;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pollen major allergen 1-1.
 OS Juniperus virginiana (Eastern red cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 NCBI_TaxID=39584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21315424; PubMed=11422137;
 RA Miodoro-Horvut T.M., Goldblum R.M., Brooks E.G.;
 RT "Identification of mutations in the genes for the pollen allergens of
 eastern red cedar (Juniperus virginiana).";
 RL Clin. Exp. Allergy 31:771-778(2001).
 DR EMBL; AF151429; AAF80166.1; -;
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;
 Query Match 76.5%; Score 62; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RPLMIIFSGNNIKL 15
 ID : ||||| |||||
 Db 87 KTLMIIFSGNNIKL 101
 RESULT 4
 Q9M4S6 PRELIMINARY; PRT; 367 AA.
 ID Q9M4S6;
 AC Q9M4S6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.

GN CUP51.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monnaie R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257491; AAF72625.1; -;
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39835 MW; B4B3C60108C2CA3 CRC64;
 Query Match 76.5%; Score 62; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RPLMIIFSGNNIKL 15
 ID : ||||| |||||
 Db 87 KALMIIFSGNNIKL 101
 RESULT 5
 Q9M4S5 PRELIMINARY; PRT; 367 AA.
 ID Q9M4S5;
 AC Q9M4S5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUP51.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257492; AAF72626.1; -;
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0E3263B741 CRC64;
 Query Match 76.5%; Score 62; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RPLMIIFSGNNIKL 15
 ID : ||||| |||||
 Db 87 KALMIIFSGNNIKL 101
 RESULT 6
 Q9M4S4 PRELIMINARY; PRT; 367 AA.
 ID Q9M4S4;
 AC Q9M4S4;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUP51.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 NCBI_TaxID=13469;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Monesalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257493; AAF72627.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 76.5%; Score 62; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPLWIFSGNNMNTKL 15
 : ||||| |||||
 Db 87 KALWIFSGNNMNTKL 101

RESULT 7
 ID Q9M4S3 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S3:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSL.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monesalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257494; AAF72628.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39832 MW; B5DFBF5A61C07A53 CRC64;

Query Match 76.5%; Score 62; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPLWIFSGNNMNTKL 15
 : ||||| |||||
 Db 87 KALWIFSGNNMNTKL 101

RESULT 8
 ID Q9M4S2 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S2:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSL.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monesalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257495; AAF72629.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39819 MW; AE7E055A61C07A53 CRC64;

Query Match 76.5%; Score 62; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPLWIFSGNNMNTKL 15
 : ||||| |||||
 Db 87 KALWIFSGNNMNTKL 101

RESULT 9
 ID Q93X51 PRELIMINARY; PRT; 367 AA.
 AC Q93X51:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Putative allergen jun o 1.
 GN JUN O 1.
 OS Juniperus oxycedrus (Prickly juniper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxID=69008;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISUE=POLLEN;
 RA TACOVACCT P., Di Felice G., Pini C.;
 RT "Cloning of Juniperus oxycedrus major allergen."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ293767; CAC48400.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase; 1.
 SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBFD1B9D7 CRC64;

Query Match 76.5%; Score 62; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPLWIFSGNNMNTKL 15
 : ||||| |||||
 Db 87 KALWIFSGNNMNTKL 101

RESULT 10
 ID Q93XL6 PRELIMINARY; PRT; 367 AA.
 AC Q93XL6:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative allergen Cup a 1 precursor.
 GN CUP A 1.
 OS Cupressus arizonica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=49011;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISUE=POLLEN;
 RA BUTERONI C., Di Felice G., Pini C.;
 RT "Cloning of Cupressus arizonica major allergen."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ278498; CAC37790.2; -
 DR Signal.
 KW SIGNAL.
 FT CHAIN
 SQ SEQUENCE 367 AA; 39809 MW; AFE97260423A9F28 CRC64;

Query Match 76.5%; Score 62; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLMIFSGNNIKL 15
 :|||||:
 Db 87 KALMIFSGNNIKL 101

RESULT 11
 ID 094LR5 PRELIMINARY; PRT; 453 AA.
 AC 094LR5;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Putative peccate lyase.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eubacteriales; Oryzaeae; Oryza.
 NCBI_TaxID=4530;

RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CV, NIPPONBARE;
 RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
 RA Brenner M., Burgess S., Hance M., Shvartbeyn M., Taitlin T.,
 RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.E.,
 RA Uterback T.R., Feldblum T.V., Quackenbush J., Salzberg S.L.,
 RA White O., Fraser C.M.;
 RL "Oryza sativa chromosome 10 BAC OSCNB0011A08 genomic sequence."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC034258; AAK54283.1;
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR KW Lyase.

SO SEQUENCE 453 AA; 48006 MW; 1411BBB1A40901DC CRC64;

Query Match 71.6%; Score 58; DB 10; Length 453;
 Best Local Similarity 64.3%; Pred. No. 0.073;
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PLMIFSGNNIKL 15
 :|||||:
 Db 140 PLMIFVAGDMTRL 153

RESULT 12
 ID 093Z04 PRELIMINARY; PRT; 501 AA.
 AC 093Z04;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE AT3G54920/F28P10_100 (Putative peccate lyase).
 GN AT3G54920.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;

RN [1]
 RC SEQUENCE FROM N.A.
 RA Shin P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Ban J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onda C.S., Palm C.J.,
 RA Phaf P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones."

Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Ban J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Tang C., Goldsmith A.D., Lee J.M., Onda C.S., Quach H.L.,
 RA Deng C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Kamaya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shin P., Southwick A., Shinzaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene At3G54920 (GI:15233132)."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY058870; AAL24257.1;
 DR EMBL: AY074331; AAL67027.1;
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR KW Lyase.

SO SEQUENCE 501 AA; 53929 MW; 9E0DE36DEF4C7AAB CRC64;

Query Match 70.4%; Score 57; DB 10; Length 501;
 Best Local Similarity 71.4%; Pred. No. 0.12;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PLMIFSGNNIKL 15
 :|||||:
 Db 146 PLMIFSSNNIKL 159

RESULT 13
 ID 09SV40 PRELIMINARY; PRT; 542 AA.
 AC 09SV40;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Peccate lyase-like protein.
 GN F28P10.100.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Choigne N., Robert C., Brothier P., Wincker P., Catolico L.,
 RA Artiguenave F., Saurin W., Weisenbach J., Mewes H.W., Mayer K.F.X.,
 RA Lemcke K., Schueller C., Quetier P., Salanoubat M.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RA BU Arabidopsis sequencing project;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL049655; CAB41092.1;
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR PRINTS: PR00807; AMBALERGEN.
 DR KW Lyase.

SO SEQUENCE 542 AA; 58573 MW; EB3779D613B72347 CRC64;

Query Match 70.4%; Score 57; DB 10; Length 542;
 Best Local Similarity 71.4%; Pred. No. 0.13;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PLMIFSGNNIKL 15
 :|||||:
 Db 146 PLMIFSSNNIKL 159

RESULT 14
 ID 064510 PRELIMINARY; PRT; 455 AA.
 AC 064510;
 RT "Arabidopsis cDNA clones."

DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Putative peccate lyase.
 GN AT2G02720.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:761-768(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC002521; AAC05350.1; -
 DR InterPro; IPR002022; Amb.allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KW lyase.
 SQ SEQUENCE 455 AA; 51257 MW; 7C6A35A767A30CA4 CRC64;
 QY 1 RPLWIFSGNNMIKL 15
 DB 166 RPLWIFSGNNMIKL 180

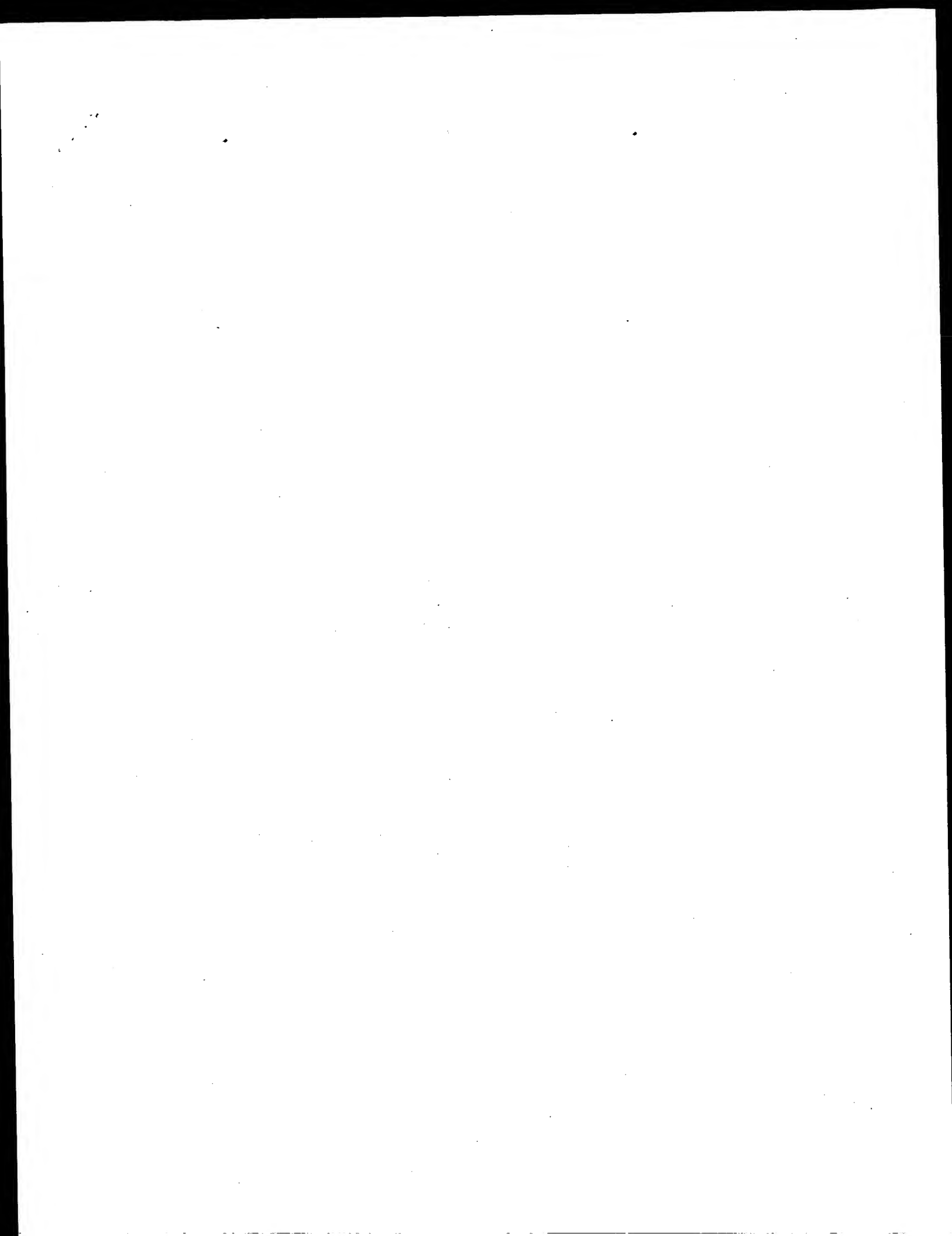
Query Match 69.1%; Score 56; DB 10; Length 455;
 Best Local Similarity 73.3%; Pred. No. 0.16; 2, Indels 0; Gaps 0;
 Matches 11, Conservative 2, Mismatches 2, Indels 0, Gaps 0;

RESULT 15
 O9SCP2 PRELIMINARY; PRT; 463 AA.
 AC O9SCP2;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE Pectate lyase-like protein.
 GN TAD2.120.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nyakatura G., Fartmann B., Dauner D., Sterr W., Holland R.,
 RA Weichselgartner M., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,
 RA Salanoubat M.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL132958; CAB64222.1; -
 DR InterPro; IPR002022; Amb.allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.

KW Lyase.
 SQ SEQUENCE 463 AA; 51225 MW; 3424B0DE4ADBA604 CRC64;
 QY 2 PLWIFPSNNMIKL 15
 DB 120 PLWIFPSNNMIKL 133

Query Match 69.1%; Score 56; DB 10; Length 463;
 Best Local Similarity 71.4%; Pred. No. 0.17;
 Matches 10, Conservative 1, Mismatches 3, Indels 0; Gaps 0;

Search completed: April 20, 2003, 13:12:59
 Job time : 14.6711 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 seconds
(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524D-29

Perfect score: 79

Sequence: 1 IFSGNNNKKMKMPY 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 73: *
2: PIR 2: *
3: PIR 3: *
4: PIR 4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	374	2	JC2124
2	79	100.0	374	2	JC2123
3	42	53.2	310	2	T29355
4	40	50.6	245	2	T17617
5	40	50.6	262	2	T43753
6	40	50.6	341	2	H86525
7	40	50.6	341	2	H72097
8	40	50.6	376	1	D89772
9	39	49.4	376	1	C64135
10	39	49.4	765	2	S72278
11	39	49.4	794	2	C86057
12	39	49.4	794	2	A98211
13	38	48.1	283	2	A99294
14	38	48.1	347	2	A99989
15	38	48.1	488	2	C84752
16	38	48.1	757	2	T16609
17	38	48.1	997	2	A60776
18	38	48.1	1344	2	T34188
19	37	46.8	81	2	T44075
20	37	46.8	238	2	E71084
21	37	46.8	275	2	F75102
22	37	46.8	376	2	G84950
23	37	46.8	492	2	A83539
24	37	46.8	522	2	ISBYSS
25	37	46.8	523	2	B85358
26	37	46.8	542	2	T66728
27	37	46.8	604	2	ACT1083
28	37	46.8	614	2	B86461
29	37	46.8	621	2	F86533

30	37	46.8	621	2	A72091
31	37	46.8	621	2	D81578
32	37	46.8	684	2	T01267
33	37	46.8	768	2	B97725
34	37	46.8	818	2	T08823
35	37	46.8	901	2	F83781
36	37	46.8	928	2	S50578
37	37	46.8	990	2	S23416
38	37	46.8	1225	2	T39255
39	37	46.8	4344	1	A53489
40	37	46.8	4639	1	A54794
41	36	45.6	122	2	A83756
42	36	45.6	161	2	C71044
43	36	45.6	174	2	T22247
44	36	45.6	189	2	T22679
45	36	45.6	233	2	H72648

ALIGNMENTS

RESULT 1

JC2124
major allergen Cry j I precursor (clone PCCI-15) - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C/Accession: JC2124

R/Score, T.: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japan

A/Reference number: JC2123; PMID:94183234; PMID:8135802

A/Accession: JC2124

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-References: GB:D26545; NID:9493633; PIDN:BAA05543.1; PID:9493634

A/Experimental source: pollen

A/Note: the authors described carbohydrate binding site for residue 279

C/Superfamily: pectate lyase LPT59

C/Keywords: glycoprotein, pollen

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-374/Product: major allergen Cry j I (clone PCCI-15) #status predicted <MAT>

F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 100.0%; Score 79; DB 2; Length 374;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFSGNNNKKMKMPY 15

Db 92 IFSGNNNKKMKMPY 106

RESULT 2

JC2123
major allergen Cry j I precursor (clone PCCI-2-2) - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C/Accession: JC2123; PC2065

R/Score, T.: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japan

A/Reference number: JC2123; PMID:94183234; PMID:8135802

A/Accession: JC2123

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-References: GB:D26544; NID:9493631; PIDN:BAA05542.1; PID:9493632

A/Experimental source: pollen

A/Accession: PC2065

A/Molecule type: protein

A/Residues: 22-53;58-81;219-232;236-258;299-307;346-372 <SO2>

C/Note: the authors described carbohydrate binding site for residue 279

C/Superfamily: pectate lyase LPT59

C/Keywords: glycoprotein, pollen

A:Experimental source: strain CML029

C:Genetics:

A:Gene: dppF_1

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match

Best Local Similarity 50.6%; Score 40; DB 2; Length 341;
Pred. No. 16;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IFSGNNNKKMKMPY 14

Db 296 ILSGNINLPRKSPM 309

RESULT 8

Db9772

hypothetical protein SA016 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: D89772

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc

ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: D89772

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-578 <KUR>

A:Cross-references: GB:BA000018; PID:g13700036; PIDN:BA841335.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA016

Query Match

Best Local Similarity 50.6%; Score 40; DB 2; Length 578;
Pred. No. 28;

Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IFSGNNNKKMKMPY 15

Db 264 VRSKALNITIKLPIH 278

RESULT 9

C64135

ribonucleoside-diphosphate reductase (EC 1.17.4.1) beta chain - Haemophilus influenzae

C:Species: Haemophilus influenzae

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: C64135

R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Welman, J.

; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Grehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: C64135

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-376 <RTGR>

A:Cross-references: GB:032839; GB:I42023; NID:g1574508; PIDN:AA23306.1; PID:g1574510; T

C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase small chain

C:Keywords: DNA replication; iron; metalloprotein; oxidoreductase

F:85,116,119,205,239,242/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, Glu, His) #str

Query Match

Best Local Similarity 49.4%; Score 39; DB 1; Length 376;
Pred. No. 27;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 FSGNNNKKMKMPY 15

Db 6 FSGNNNKKMKMPY 19

RESULT 10

S72278

ATP-dependent Clp proteinase (EC 3.4.21.-) homolog - Plasmodium falciparum

C:Species: Plasmodium falciparum

C>Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 20-Jun-2000

C:Accession: S72278; S78483

R:Wilson, R.J.M.; Denny, P.W.; Preiser, P.R.; Rangachari, K.; Roberts, K.; Roy, A.; W

J. Mol. Biol. 261, 155-172, 1996

A>Title: Complete gene map of the plasmodium-like DNA of the malaria parasite Plasmodium

A:Reference number: S72277; MUID:96346169; PMID:8757284

A:Accession: S72278

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-765 <WIL>

A:Cross-references: EMBL:X95276

R:Wilson, R.J.M.

submitted to the EMBL Data Library, January 1996

A:Reference number: S78483

A:Accession: S78483

A:Molecule type: DNA

A:Residues: 1-95, 'N', '96-765 <WIL>

A:Cross-references: EMBL:X95276; NID:g1171591; PIDN:CAA64596.1; PID:g1171612

C:Genetics:

A:Gene: clpC

A:Note: this apparently degenerate plasmid is referred to as the apicoplast

C:Superfamily: endopeptidase Clp ATP-binding chain

C:Keywords: hydrolase; plasmid; serine proteinase

Query Match

Best Local Similarity 49.4%; Score 39; DB 2; Length 765;
Pred. No. 59;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IFSGNNNKKMKMPY 11

Db 143 IFSGNNNKKMKMPY 153

RESULT 11

C86057

hypothetical protein Z5214 [imported] - Escherichia coli (strain O157:H7, substrain EF

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: C86057

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayh

iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; DiMaio, E.; Potamousis, K.; Apodac

Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A65480; MUID:21074935; PMID:11206551

A:Accession: C86057

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-794 <STO>

A:Cross-references: GB:AE005174; NID:g12518567; PIDN:AGS8919.1; GSPDB:GN00145; UWGP:Z

A:Experimental source: strain O157:H7, substrain EDJ933

C:Genetics:

A:Gene: Z5214

Query Match

Best Local Similarity 49.4%; Score 39; DB 2; Length 794;
Pred. No. 61;

Matches 8; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

Qy 1 IFSGNNNKKMKMPY 15

Db 456 ILSGNLANVQKIPY 472

RESULT 12

A98211

hypothetical protein Eca657 [imported] - Escherichia coli (strain O157:H7, substrain R

C:Species: Escherichia coli

C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C/Accession: A98211
 R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gatawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
 A/Reference number: A99629; PMID:21156231; PMID:11258796
 A/Accession: A98211
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-794 <HAY>
 A/Cross-references: GB:BA000007; PIDN:BA038080.1; PID:g13364132; GSPDB:GN00154
 A/Experimental source: strain O157:H7, substrain RMD 0509952
 C/Genetics:
 A/Gene: ECG4657

Query Match 49.4%; Score 39; DB 2; Length 794;
 Best Local Similarity 47.1%; Pred. No. 61;
 Matches 8; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

OY 1 IFSGNM--NIKMKPMY 15
 |||||:|:|:|:|:
 Db 456 ILSGNLAANTQVKIPVY 472

RESULT 13
 A99294
 hypothetical protein SSO1374 [imported] - *Sulfolobus solfataricus*

C/Species: *Sulfolobus solfataricus*
 C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C/Accession: A99294
 R/Shi, Q.; Singh, R.K.; Cafaloni, F.; Zivanovic, Y.; Allard, G.; Anayez, M.J.; Chan-
 Jiong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A/Description: *Sulfolobus solfataricus* complete genome.
 A/Reference number: A99139
 A/Accession: A99294
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-283 <KUR>
 A/Cross-references: GB:AE006641; NID:g13814586; PIDN:AAK41608.1; GSPDB:GN00155
 C/Genetics:
 A/Gene: SSO1374

Query Match 48.1%; Score 38; DB 2; Length 283;
 Best Local Similarity 38.5%; Pred. No. 30;
 Matches 5; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

OY 3 SGNNNIKMKPMY 15
 |||||:|:|:|:|:
 Db 159 AGNTSIPKIPHL 171

RESULT 14

A99989
 cyclin B [imported] - *Galliardia theta nucleomorph*

C/Species: *nucleomorph Galliardia theta*
 A/Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C/Accession: A99989
 R/Douglas, S.; Zanner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei-

Nature 410, 1091-1096, 2001
 A/Title: The highly reduced genome of an enslaved algal nucleus.
 A/Reference number: A99082; PMID:11323671; PMID:11323671

A/Accession: A99989
 A/Status: preliminary
 A/Molecule type: DNA

A/Residues: 1-347 <DOU>
 A/Cross-references: GB:AF165818; NID:g13794469; PIDN:AAK39844.1; GSPDB:GN00150
 C/Genetics:
 A/Gene: cycB
 A/Map position: 1

A/Genome: nucleomorph
 C/Keywords: nucleomorph

Query Match 48.1%; Score 38; DB 2; Length 347;
 Best Local Similarity 42.9%; Pred. No. 37;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 2 FSGNNNIKMKPMY 15
 |||||:|:|:|:|:
 Db 279 FSGNISIKITMNI 292

RESULT 15

C84752
 hypothetical protein At2g34090 [imported] - *Arabidopsis thaliana*

C/Species: *Arabidopsis thaliana* (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C/Accession: C84752

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A/Reference number: A84420; PMID:20083487; PMID:10617197

A/Accession: C84752

A/Status: preliminary
 A/Molecule type: DNA

A/Residues: 1488 <STO>
 A/Cross-references: GB:AE002093; NID:g2342730; PIDN:AAE67628.1; GSPDB:GN00139

C/Genetics:
 A/Gene: At2g34090

A/Map position: 2
 C/Superfamily: Saccharomyces cerevisiae probable membrane protein YNL011c

Query Match 48.1%; Score 38; DB 2; Length 488;
 Best Local Similarity 58.3%; Pred. No. 54;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 GNNNIKMKPMY 15
 |||||:|:|:|:|:
 Db 380 GDTNIRLKNPT 391

Search completed: April 20, 2003, 13:15:40
 Job time: 8.07895 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524d-29
Perfect score: 79
Sequence: 1 IFSGNNMIKMKPMK 15

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	374	1 SBP_CRYJA	P18632 cryptomeria
2	68	86.1	367	1 MPAL_JUNAS	P81294 juniperus a
3	64	81.0	346	1 MPAL_CUPAR	O96c99 cupressus a
4	60	75.9	375	1 MPAL_CHABO	O96385 chamaecypar
5	42	53.2	283	1 PORI_CABEL	Q21752 caenohabdi
6	39	49.4	375	1 RIR2_HABIN	P43755 haemophilus
7	38	48.1	111	1 FER_BUCAP	O51882 buchnera ap
8	37	46.8	376	1 RIR2_BUCAI	P17967 saccharomyc
9	37	46.8	522	1 PDI_YEAST	O9k275 chlamydia p
10	37	46.8	621	1 Y344_CHLPN	P40048 saccharomyc
11	37	46.8	928	1 PTP3_YEAST	P20195 ataphylococ
12	37	46.8	986	1 EP1B_STREP	P45444 emeticella
13	37	46.8	4344	1 DYHC_EMENT	P37276 drosophila
14	37	46.8	4639	1 DYHC_DROME	O99394 saccharomyc
15	36	45.6	268	1 TRJ3_YEAST	O9hkd2 thermoplas
16	36	45.6	424	1 TRPB_THRAC	Q03210 saccharomyc
17	36	45.6	505	1 YML8_YEAST	P42433 bacillus su
18	36	45.6	770	1 NASE_BACSU	P45998 haemophilus
19	36	45.6	837	1 HFC3_HABIN	P40559 saccharomyc
20	36	45.6	946	1 YIA2_YEAST	P28692 saccharomyc
21	36	45.6	1080	1 NRK1_YEAST	P28692 saccharomyc
22	36	45.6	1286	1 YKVS_YEAST	P28692 saccharomyc
23	36	45.6	1536	1 Y984_THEMA	O9x079 thermotoga
24	36	45.6	4349	1 DYHC_FUSSO	P78716 fusarium so
25	35	44.3	327	1 SRP8_SCHPO	O13950 schizosach
26	35	44.3	367	1 YC36_HABIN	P44132 haemophilus
27	35	44.3	492	1 SYNM_YEAST	P25345 saccharomyc
28	35	44.3	555	1 POLG_DEN1T	P27909 dengue viru
29	35	44.3	558	1 POLG_DEN1T	P27909 dengue viru
30	35	44.3	559	1 SYD_MYCPN	P75068 mycoplasma
31	35	44.3	581	1 F1B_ADENO	P17818 human adeno
32	35	44.3	791	1 POLG_DEN1C	P27910 dengue viru
33	35	44.3	792	1 POLG_DEN1B	P27910 dengue viru

ALIGNMENTS

RESULT 1	ID	SBP_CRYJA	STANDARD;	PRT;	374 AA.
AC	P18632;				
DT	01-NOV-1990 (Rel. 16, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	Sugi basic protein precursor (SBP) (Major allergen Cry j 1) (Cry j 1).				
OS	Cryptomeria japonica (Japanese cedar).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.				
OK	NCBI_Taxid=3369;				
RN	(1)				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE=Pollen;				
RA	MEDLINE=94183234; PubMed=8135802;				
RX	Sone T., Komiya N., Shimizu K., Kusakabe T., Morikubo K.,				
RA	kind K.,				
RT	"Cloning and sequencing of cDNA coding for Cry j I, a major allergen				
RT	of Japanese cedar pollen."				
RL	Biochem. Biophys. Res. Commun. 199:619-625 (1994).				
RN	(2)				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Pollen;				
RA	Namua M., Kurose M., Torioge K., Fukuda S., Kurimoto M.,				
RN	Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.				
RP	SEQUENCE OF 22-41.				
RC	TISSUE=Pollen;				
RX	MEDLINE=89031257; PubMed=3181436;				
RA	Tanai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inoue S.,				
RA	Matuhashi T.,				
RT	"N-terminal amino acid sequence of a major allergen of Japanese cedar				
RT	pollen (Cry j I)."				
RL	FEBS Lett. 239:329-332 (1988).				
RN	(4)				
RP	CARBOHYDRATES.				
RC	TISSUE=Pollen;				
RX	MEDLINE=95003748; PubMed=7920021;				
RA	Hijikata A., Matsumoto I., Kojima K., Ogawa H.,				
RT	"Antigenicity of the oligosaccharide moiety of the Japanese cedar				
RT	(Cryptomeria japonica) pollen allergen, Cry j I."				
RL	Int. Arch. Allergy Immunol. 105:198-202 (1994).				
RN	(5)				
RP	STRUCTURE OF CARBOHYDRATES.				
RC	TISSUE=Pollen;				
RX	MEDLINE=95332249; PubMed=7608114;				
RA	Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Usui M.,				
RA	Fukuda S., Hanzawa H., Hanayama H., Kurimoto M.,				
RT	"Carbohydrate structures of the glycoprotein allergen Cry j I from				
RT	Japanese cedar (Cryptomeria japonica) pollen."				
RL	J. Biochem. 117:289-295 (1995).				
CC	-1- PWM: CONTAINS FUCCSE/YLCCSE-CONTAINING N-LINKED OLIGOSACCHARIDES.				
CC	-1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR				
CC	POLLEN, THE MOST COMMON POLLEN ALLERGEN IN JAPAN.				
CC	-1- MISCELLANEOUS: THE SEQUENCE OF CRY J I FROM A IS SHOWN HERE. FORM				

34	35	44.3	792	1	POLG_DEN1A	P27912 dengue viru
35	35	44.3	959	1	VILI_DICDI	P36418 dictyostell
36	35	44.3	1226	1	POLG_DEN1W	P17763 dengue viru
37	35	44.3	1462	1	NC02_MOUSE	O61026 dengue viru
38	35	44.3	1464	1	NC02_HUMAN	O15596 homo sapien
39	35	44.3	1477	1	AL13_RAT	P14046 rattus norv
40	35	44.3	1541	1	NCX3_HUMAN	O9740 homo sapien
41	35	44.3	2258	1	PAS_FIG	O991P1 sus scrofa
42	35	44.3	3390	1	POLG_DEN3	P27915 d genome po
43	35	44.3	3396	1	POLG_DEN1S	P33478 d genome po
44	34	43.0	130	1	MTA0_YEAST	P09091 saccharomyc
45	34	43.0	208	1	C130_ARATH	G91q17 arabidopsis

KRP
 SEQUENCE FROM N.A.
 STRAIN-Rd / KM20 / ATCC 51907;
 MEDLINE=95350630, PubMed=7542800.
 Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Lera
 kerlaevge A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Ullrich T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geiselman N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512 (1995).
 CC -1- FUNCTION: CATALYZES THE BIOSYNTHESIS OF DEOXYRIBONUCLEOTIDES FROM
 CC THE CORRESPONDING RIBONUCLEOTIDES, PRECURSORS THAT ARE NECESSARY
 CC FOR DNA SYNTHESIS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
 CC thioedoxin + H(2)O = ribonucleoside diphosphate + reduced
 CC thioedoxin.
 CC -1- COFACTOR: BINDS 2 IRON IONS (BY SIMILARITY).
 CC -1- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
 CC -1- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
 CC SMALL CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; U32839; AAC23306.1; -;
 CC DR HSSP; P00453; IXIX.
 CC DR TIGR; H11600; -;
 CC DR InterPro; IPR000358; RibonucL_redctase.
 CC DR Pfam; PF00268; ribonuc_red_sm; 1.
 CC DR PROSITE; PS00368; RIBORED_SMALL; 1.
 CC KW Oxidoreductase; DNA replication; Iron; Complete proteome.
 CC FT INTL MET 0
 CC FT METAL 84
 CC FT METAL 115
 CC FT METAL 118
 CC FT METAL 118
 CC FT METAL 204
 CC FT METAL 238
 CC FT METAL 241
 CC FT ACT SITE 122
 CC SQ SEQUENCE 375 AA; 43194 MW; 7BA30D8C96A5B35 CRC64;
 CC
 CC Query Match 49.4%; Score 39; DB 1; Length 375;
 CC Best Local Similarity 57.1%; Pred. No. 6.1;
 CC Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

RT "Sequence analysis of a 34.7-kb DNA segment from the genome of
 RT Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA,
 RT the acp operon, gldA, and rho.";
 RL Curr. Microbiol. 36:158-163 (1998).
 CC -1- FUNCTION: FERREDOXIN ARE IRON-SULFUR PROTEINS THAT TRANSFER
 CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
 CC -1- COFACTOR: BINDS 1 2FE-2S CLUSTER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ADRENODOXIN / PUTIDREDOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF008210; AAC38120.1; -;
 CC DR HSSP; P00257; 1E6E.
 CC DR InterPro; IPR001055; Adrenodoxin.
 CC DR Pfam; PF00111; fer2; 1.
 CC DR PRINTS; PR00355; ADRENODOXIN.
 CC DR PROSITE; PS00814; ADX; 1.
 CC KW Electron transport; Iron-sulfur.
 CC FT METAL 42
 CC FT METAL 42
 CC FT METAL 48
 CC FT METAL 51
 CC FT METAL 87
 CC SQ SEQUENCE 111 AA; 12458 MW; A0DB3A1209A935D CRC64;
 CC
 CC Query Match 48.1%; Score 38; DB 1; Length 111;
 CC Best Local Similarity 41.7%; Pred. No. 2.3;
 CC Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 4 GNNNTIKLKMPY 15
 DB 92 GNIDIEVOIPLY 103

RESULT 8
 R122_BUCAI STANDARD; PRT; 376 AA.
 AC P57275;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ribonucleoside diphosphate reductase beta chain (EC 1.17.4.1)
 DE (Ribonucleoside diphosphate small subunit).
 GN NRDB OR BU178.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tokyo 1998;
 RX MEDLINE=20445173; PubMed=10933077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RT Buchnera sp. APS.";
 RL Nature 407:81-86 (2000).
 CC -1- FUNCTION: CATALYZES THE BIOSYNTHESIS OF DEOXYRIBONUCLEOTIDES FROM
 CC THE CORRESPONDING RIBONUCLEOTIDES, PRECURSORS THAT ARE NECESSARY
 CC FOR DNA SYNTHESIS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
 CC thioedoxin + H(2)O = ribonucleoside diphosphate + reduced
 CC thioedoxin.
 CC -1- COFACTOR: BINDS 2 IRON IONS (BY SIMILARITY).
 CC -1- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
 CC -1- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE

CC SMALL CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AP001118; BAB12895.1; -

DR HSSP; P00453; 1XIK.

DR InterPro: IPR000358; Ribonuc[redacted] redctase.

DR Pfam; PF00268; ribonuc_red_sm; 1.

DR PROSITE; PS00368; RIBONUC_SMALL; 1.

KW Oxidoreductase; DNA replication; Iron; Complete proteome.

FT METAL 85 85 IRON 1 (BY SIMILARITY).

FT METAL 116 116 IRON 1 AND 2 (BY SIMILARITY).

FT METAL 119 119 IRON 1 (BY SIMILARITY).

FT METAL 205 205 IRON 2 (BY SIMILARITY).

FT METAL 239 239 IRON 2 (BY SIMILARITY).

FT METAL 242 242 IRON 2 (BY SIMILARITY).

FT ACT_SITE 123 123 BY SIMILARITY.

SQ SEQUENCE 376 AA; 44145 MW; BECC1B02F77CBB32 CRC64;

Query Match 46.8%; Score 37; DB 1; Length 376;

Best Local Similarity 53.3%; Pred. No. 15;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 IFSGNMNKLKMPY 15

Db 5 IFSKKNNQKKEPMF 19

RESULT 9

PDI_YEAST

ID PDI_YEAST STANDARD; PRT; 522 AA.

AC P17967;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Protein disulfide isomerase precursor (PDI) (EC 5.3.4.1) (Thioredoxin-

DE related glycoprotein 1).

GN PDI1 OR MFP1 OR TRG1 OR YCL043C OR YCL43C OR YCL313.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN=X2180-1A, and TMS;

RX MEDLINE=92105067; PubMed=1761527;

RA Tachikawa H., Miura T., Katakura Y., Mizunaga T.;

RT "Molecular structure of a yeast gene, PDI1, encoding protein

RT disulfide isomerase that is essential for cell growth.";

RL J. Biochem. 110:306-313(1991).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=91289690; PubMed=2063627;

RA Scherens B., Dubois E., Messenguy F.;

RT "Determination of the sequence of the yeast YCL313 gene localized on

RT chromosome III. Homology with the protein disulfide isomerase (PDI

RL yeast product) of other organisms.";

RL Yeast 7:185-193(1991).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=912359586; PubMed=1840656;

RA Lamantia M., Miura T., Tachikawa H., Kaplan H.A., Lennarz W.J.;

RA Mizunaga T.;

RT "Glycosylation site binding protein and protein disulfide isomerase

RT are identical and essential for cell viability in yeast.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:4453-4457(1991).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=92105117; PubMed=1761554;

RA Gauthier R., Brazeau C., Janetzky B., Foerster H.H., Ebhrecht I.M.;

RA Lehle L., Kuentzel H.;

RT "The Saccharomyces cerevisiae TRG1 gene is essential for growth and

RT encodes a luminal endoplasmic reticulum glycoprotein involved in the

RT maturation of vacuolar carboxypeptidase.";

RL U. Biol. Chem. 266:24557-24563(1991).

RN [5]

RP SEQUENCE FROM N.A.

RX MEDLINE=92397595; PubMed=1523890;

RA Scherens B., Messenguy F., Gicot D., Dubois E.;

RT "The complete sequence of a 9,343 bp segment on the left arm of

RT chromosome III reveals five open reading frames including glucokinase

RL and the protein disulfide isomerase.";

RL Yeast 8:577-586(1992).

RN [6]

RP SEQUENCE FROM N.A.

RX MEDLINE=92104510; PubMed=1761235;

RA Fargnhar R., Honey N., Murant S.J., Boesier P., Schultz L.;

RA Montogomery D., Ellis R.W., Freedman R.B., Tuile M.F.;

RT "Protein disulfide isomerase is essential for viability in

RT Saccharomyces cerevisiae.";

RL Gene 108:81-89(1991).

CC -1- FUNCTION: PARTICIPATES IN THE FOLDING OF PROTEINS CONTAINING

CC DISULFIDE BONDS, MAY BE INVOLVED IN GLYCOSYLATION, PROLYL

CC HYDROXYMATION AND TRIGLYCERIDE TRANSFER.

CC -1- CATALYTIC ACTIVITY: Rearrangement of both intrachain and

CC interchain disulfide bonds in proteins to form the native

CC structures.

CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (potential).

CC -1- PTM: The N-terminus is blocked.

CC -1- SIMILARITY: CONTAINS 2 THIOREDOXIN DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X52313; CAA36550.1; -

DR EMBL; X57712; CAA40883.1; -

DR EMBL; M62815; AAA34948.1; -

DR EMBL; M76882; AAA35169.1; -

DR EMBL; X59720; CAA42373.1; -

DR EMBL; X54535; CAA38402.1; -

DR EMBL; D00842; BAA00723.1; -

DR PIR; JX0182; ISBYSS.

DR PIR; S19372; S19372.

DR PIR; S15050; S15050.

DR PIR; S25349; S25349.

DR PIR; A41713; A41713.

DR PIR; U50634; U50634.

DR HSSP; P07237; IMEX.

DR SGD; S0000548; PDI1.

DR InterPro: IPR000886; ER target.

DR InterPro: IPR000663; Thioredo.

DR Pfam; PF00085; thioredo_2.

DR PRINTS; PR00421; THIREDOXIN.

DR TIGRFAMs; TIGR01126; pdi_dom; 2.

DR TIGRFAMs; TIGR01130; ER_PDI_fam; 1.

DR PROSITE; PS00014; ER_TARGET_1.

DR PROSITE; PS00194; THIREDOXIN_2.

KW Redox-active center; Isomerase; Endoplasmic reticulum; Repeat; Signal.

FT SIGNAL 1 28

FT CHAIN 1 522

FT DISULFID 61 522

FT DISULFID 64 64

FT CARBOHYD 406 409

FT CARBOHYD 82 82

FT CARBOHYD 117 117

FT CARBOHYD 155 155

FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 519 522 PREVENT SECRETION FROM ER.
FT CONFLICT 33 52 AVALKALDSNFTSRTW (IN REF. 3).
FT CONFLICT 83 83 I -> V (IN REF. 4).
FT CONFLICT 114 114 S -> R (IN REF. 6).
FT CONFLICT 143 143 V -> S (IN REF. 4).
FT CONFLICT 146 146 MISSING (IN REF. 4).
FT CONFLICT 168 168 K -> E (IN REF. 3).
FT CONFLICT 197 197 D -> R (IN REF. 3).
FT CONFLICT 215 215 A -> V (IN REF. 1; AA SEQUENCE).
FT CONFLICT 226 226 A -> V (IN REF. 4).
FT CONFLICT 333 333 E -> S (IN REF. 1; AA SEQUENCE).
FT CONFLICT 351 351 S -> P (IN REF. 3).
FT CONFLICT 455 455 L -> F (IN REF. 3).
FT CONFLICT 458 458 G -> S (IN REF. 4).
FT CONFLICT 505 505 A -> AADMAEA (IN REF. 4 AND 6).
SQ SEQUENCE 522 AA; 58227 MW; 69CF3B05D7F74C94 CRC64;

Query Match 46.8%; Score 37; DB 1; Length 522;
Best Local Similarity 38.5%; Pred. No. 22;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 3 GGNNNIKLKMPY 15
DB 302 AGNNLMKEQFPLF 314

RESULT 10
Y344 CHLPN STANDARD; PRT; 621 AA.
AC Q9K275; Q9JSG0; Q9Z8U9;
ID 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical zinc metalloprotease (CP0344/CP0344 (EC 3.4.24.-)).
GN CP0344 OR CP0416 OR CP0344.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=9920606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammell C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Ulfenback T., Berry K., Bass S.,
RA Linher K., Weidman U., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin J., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shiba T., Ishikawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shira T., Ishi K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:12311-2314(2000).
CC -1- COFACTOR: Zinc (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

```

```

CC (By similarity). BELONGS TO PEPTIDASE FAMILY M50B.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M50B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB001618; AAD18488.1; -
DR EMBL: AB002203; AAF8260.1; -
DR EMBL: AP002546; BAA8552.1; -
DR TIGR: CP0416; -
DR InterPro: IPR001478; PDZ.
DR SMART: SM00228; PDZ; 2.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein; Hydrolase; Metalloprotease; Zinc; Transmembrane;
KW Inner membrane; Complete proteome.
KM Inner membrane; Complete proteome.
FT METAL 20 20 ZINC (CATALYTIC) (POTENTIAL).
FT ACT SITE 21 21 POTENTIAL.
FT METAL 24 24 ZINC (CATALYTIC) (POTENTIAL).
FT TRANSMEM 103 125 POTENTIAL.
FT TRANSMEM 561 583 POTENTIAL.
FT TRANSMEM 596 613 POTENTIAL.
FT CONFLICT 224 224 G -> S (IN REF. 1).
FT CONFLICT 584 584 T -> K (IN REF. 1).
FT CONFLICT 598 598 V -> D (IN REF. 3).
SQ SEQUENCE 621 AA; 69820 MW; 30B09776F8A271BF CRC64;

Query Match 46.8%; Score 37; DB 1; Length 621;
Best Local Similarity 33.3%; Pred. No. 27;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
DB 178 LLEGHNLRIKRGY 192

RESULT 11
PTP3 YEAST STANDARD; PRT; 928 AA.
AC P40048;
ID 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-FEB-1998 (Rel. 37, Last annotation update)
DE Protein-tyrosine phosphatase 3 (EC 3.1.3.48) (PTPASE 3).
GN PTP3 OR YER075C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RX MEDLINE=97367951; PubMed=9224718;
RA Zhan X.-L., Deschenes R.D., Guan K.-L.;
RT "Differential regulation of FUS3 MAP kinase by tyrosine-specific
RT phosphatases PTP2/PTP3 and dual-specificity phosphatase MSG5 in
RT Saccharomyces cerevisiae.";
RL Genes Dev. 11:1690-1702(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Huntke-Smith S.,
RA Hyman R., Kayser A., Komp C., Isakari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RT Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
RL

```

CC -!-FUNCTION: MAJOR PHOSPHATASE RESPONSIBLE FOR TYROSINE
 CC DEPHOSPHORYLATION OF MAP KINASE FUS3 TO INACTIVATE ITS ACTIVITY;
 CC IT ALSO HAS IMPORTANT ROLES, ALONG WITH MSG5, IN THE INACTIVATION
 CC OF FUS3 FOLLOWING PHEROMONE STIMULATION.
 CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate
 CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF006304; AAB70811.1; -
 CC DR EMBL; U18814; AAB64614.1; -
 CC DR HSSP; P29350; 1GWZ.
 CC DR SGD; S0000877; PTP3.
 CC DR InterPro; IPR001763; Rhodanese-like.
 CC DR InterPro; IPR000387; TYR phosphatase.
 CC DR InterPro; IPR000242; TYR_PP.
 CC DR Pfam; PF00102; Y.phosphatase; 1.
 CC DR Pfam; PF00581; Rhodanese; 1.
 CC DR SMART; SM00194; PTPC; 1.
 CC DR SMART; SM00450; RHOD; 1.
 CC DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 CC DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 CC DR PROSITE; PS50055; TYR_PHOSPHATASE_PTE; 1.
 CC DR Hydrolase.
 CC FT ACT_SITE 804 804 BY SIMILARITY.
 CC FT DOMAIN 308 311 POLY-SER.
 CC FT DOMAIN 592 597 POLY-THR.
 CC FT DOMAIN 702 715 POLY-ASN.
 CC FT DOMAIN 724 732 POLY-ASP.
 CC SEQUENCE 928 AA; 105250 MW; 6DECBA26B7ACB6B CRC64;
 CC
 CC Query Match 46.8%; Score 37; DB 1; Length 928;
 CC Best Local Similarity 63.6%; Pred. No. 43;
 CC Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 3 SGNNMIRKMP 13
 CC | : : : : :
 CC Db 272 SPQNNLTKVP 282
 CC
 CC RESULT 12
 CC EPIB_STAEP STANDARD; PRT; 986 AA.
 CC AC P30195;
 CC DT 01-APR-1993 (Rel. 25, Created)
 CC DT 01-APR-1993 (Rel. 25, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Epidermal biosynthesis protein epib.
 CC GN EPIB.
 CC OS Staphylococcus epidermidis.
 CC OG Plasmid pTU 32.
 CC OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 CC OX NCBI_TaxID=1282;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-TU 3298 / DSM 3095;
 CC RX MEDLINE=92155237; PubMed=1740156;
 CC RA Schell N., Engelse G., Augustin J., Rosenstein R., Ungermann V.,
 CC Goetz F., Entian K.-D.;
 CC RT "Analysis of genes involved in the biosynthesis of antibiotic
 CC epidermin";
 CC Eur. J. Biochem. 204:57-68(1992).
 CC -!- FUNCTION: INVOLVED IN THE POSTTRANSLATIONAL MODIFICATION OF THE
 CC LANTIBIOTIC EPIDERMIN.
 CC CC -!- SUBCELLULAR LOCATION: POSSIBLY ASSOCIATED WITH, AND ANCHORED TO,

CC THE CYTOPLASMIC SIDE OF THE MEMBRANE.
 CC -!- SIMILARITY: TO B.SUBTILIS SPAN AND L.LACTIS NISB.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X62386; CAA44253.1; -
 CC KW Plasmid.
 CC SEQUENCE 986 AA; 116708 MW; 343CE98526847850 CRC64;
 CC
 CC Query Match 46.8%; Score 37; DB 1; Length 986;
 CC Best Local Similarity 70.0%; Pred. No. 46;
 CC Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 2 FSGNNMIRK 11
 CC | : : : : :
 CC Db 462 FTGNNMIRK 471
 CC
 CC RESULT 13
 CC DYHC_EMENTI STANDARD; PRT; 4344 AA.
 CC AC P45474;
 CC DT 01-NOV-1995 (Rel. 32, Created)
 CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE Dynein heavy chain, cytosolic (DYHC).
 CC GN NUDA.
 CC OS Emericella nidulans (Aspergillus nidulans).
 CC CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 CC OC Eurotiiales; Trichocomaceae; Emericella.
 CC OX NCBI_TaxID=5072;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=94181539; PubMed=8134356;
 CC RA Xiang X., Beckwith S.M., Morris R.N.;
 CC RT "Cytoplasmic dynein is involved in nuclear migration in Aspergillus
 CC nidulans";
 CC Proc. Natl. Acad. Sci. U.S.A. 91:2100-2104(1994).
 CC -!- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
 CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
 CC ORGANELLES ALONG MICROTUBULES. REQUIRED TO MAINTAIN UNIFORM
 CC NUCLEAR DISTRIBUTION IN HYPAE.
 CC -!- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
 CC INTERMEDIATE AND LIGHT CHAINS.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; U03904; AAA18338.1; -
 CC DR PIR; A53489; A53489.
 CC DR InterPro; IPR004273; Dynein heavy.
 CC DR Pfam; PF03028; Dynein heavy; 1.
 CC KW Motor protein; Microtubules;
 CC FT DOMAIN 667 687 Dynein; ATP-binding, Coiled coil.
 CC FT 913 933 COILED COIL (POTENTIAL).
 CC FT 1321 1341 COILED COIL (POTENTIAL).
 CC FT 1548 1564 COILED COIL (POTENTIAL).
 CC FT 1627 1656 COILED COIL (POTENTIAL).
 CC FT 2035 2063 COILED COIL (POTENTIAL).
 CC FT 3179 3270 COILED COIL (POTENTIAL).
 CC DOMAIN 3179 3270 COILED COIL (POTENTIAL).

FT DOMAIN 3407 3463 COILED COIL (POTENTIAL).
 FT DOMAIN 3700 3798 COILED COIL (POTENTIAL).
 FT NP BIND 1933 1940 ATP (POTENTIAL).
 FT NP BIND 2223 2230 ATP (POTENTIAL).
 FT NP BIND 2592 2599 ATP (POTENTIAL).
 FT NP BIND 2932 2939 ATP (POTENTIAL).
 SQ SEQUENCE 4344 AA; 492470 MW; 302A588CAF8BD019 CRC64;

Query Match 46.8%; Score 37; DB 1; Length 4344;
 Best Local Similarity 42.9%; Pred. No. 2.6e+02;
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IFSGNNMIKMKPM 14
 Db 1584 VFTGNADIKILPV 1597

RESULT 14
 ID DYHC DROME STANDARD; PRT; 4639 AA.
 AC 37276;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Dynein heavy chain, cytosolic (DYHC).
 OS CDHC OR DHC64C.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=94375524; PubMed=8089180;
 RA Li M., McGrail M., Serr M., Hays T.S.;
 RT "Drosophila cytoplasmic dynein, a microtubule motor that is
 RT asymmetrically localized in the oocyte."
 RL J. Cell Biol. 126:1475-1494(1994).
 [2]
 RN SEQUENCE OF 1877-1998 FROM N.A.
 RX MEDLINE=94243034; PubMed=8186464;
 RA Rasmussen K., Serr M., Gerner J., Gibbons I., Hays T.S.;
 RT "A family of dynein genes in Drosophila melanogaster."
 RL Mol. Biol. Cell 5:45-55(1994).
 CC -1- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
 CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
 CC ORGANELLES ALONG MICROTUBULES.
 CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
 CC INTERMEDIATE AND LIGHT CHAINS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L23195; AAA60323.1; -;
 DR EMBL; L25122; AAA28482.1; -;
 DR Flybase; FBgn010349; Dhc64C.
 DR InterPro; IPR004273; Dynein_heavy.
 DR Pfam; PF03028; Dynein_heavy; 1.
 KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
 FT DOMAIN 530 565
 FT DOMAIN 774 794
 FT DOMAIN 1264 1368
 FT DOMAIN 1999 2027
 FT DOMAIN 3189 3261
 FT DOMAIN 3382 3478
 FT DOMAIN 3723 3782
 COILED COIL (POTENTIAL).
 COILED COIL (POTENTIAL).
 COILED COIL (POTENTIAL).
 COILED COIL (POTENTIAL).
 COILED COIL (POTENTIAL).
 COILED COIL (POTENTIAL).

FT NP BIND 1895 1902 ATP (POTENTIAL).
 FT NP BIND 2210 2217 ATP (POTENTIAL).
 FT NP BIND 2580 2587 ATP (POTENTIAL).
 FT NP BIND 2922 2929 ATP (POTENTIAL).
 SQ SEQUENCE 4639 AA; 530152 MW; 057A7D8800CCD07E CRC64;
 Query Match 46.8%; Score 37; DB 1; Length 4639;
 Best Local Similarity 50.0%; Pred. No. 2.8e+02;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IFSGNNMIKMKPM 14
 Db 1540 IFSGADIKILPV 1553

RESULT 15
 ID TR33 YEAST STANDARD; PRT; 268 AA.
 AC 099394;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Transport protein particle 33 kDa subunit (TRAP 33 kDa subunit).
 OS TR33 OR YOR115C OR O3251 OR YOR3251C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_Taxid=4932;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=S288C / FY1679;
 RC MEDLINE=97060020; PubMed=8904341;
 RA Wiemann S., Rechmann S., Benes V., Voss H., Schwager C., Vlcek C.,
 RA Stegemann U., Zimmermann J., Bfelle H., Paces V., Ansoerge W.;
 RT "Sequencing and analysis of 51 kb on the right arm of chromosome XV
 RT from Saccharomyces cerevisiae reveals 30 open reading frames."
 RL Yeast 12:281-288(1996).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=97344368; PubMed=9200815;
 RA Voss H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoru C.,
 RA Schwager C., Paces V., Sander C., Ansoerge W.;
 RT "DNA sequencing and analysis of 130 kb from yeast chromosome XV."
 RL Yeast 13:655-672(1997).
 CC -1- FUNCTION: TRAP plays a key role in the late stages of endoplasmic
 CC reticulum to Golgi traffic.
 CC -1- SUBUNIT: Part of multisubunit complex composed of BET3, BET5,
 CC TRS20, TRS23, TRS31, TRS33, TRS65, TRS85, TRS120 and TRS130.
 CC -1- SUBCELLULAR LOCATION: cis-Golgi complex.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X90518; CAA62124.1; -;
 DR EMBL; X94335; CAA64035.1; -;
 DR EMBL; Z75023; CAA99313.1; -;
 DR SCD; S0005641; TRS33.
 KW Transport; Endoplasmic reticulum; Golgi stack.
 SQ SEQUENCE 268 AA; 30749 MW; 248069D7A4ED363 CRC64;

Query Match 45.6%; Score 36; DB 1; Length 268;
 Best Local Similarity 63.6%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

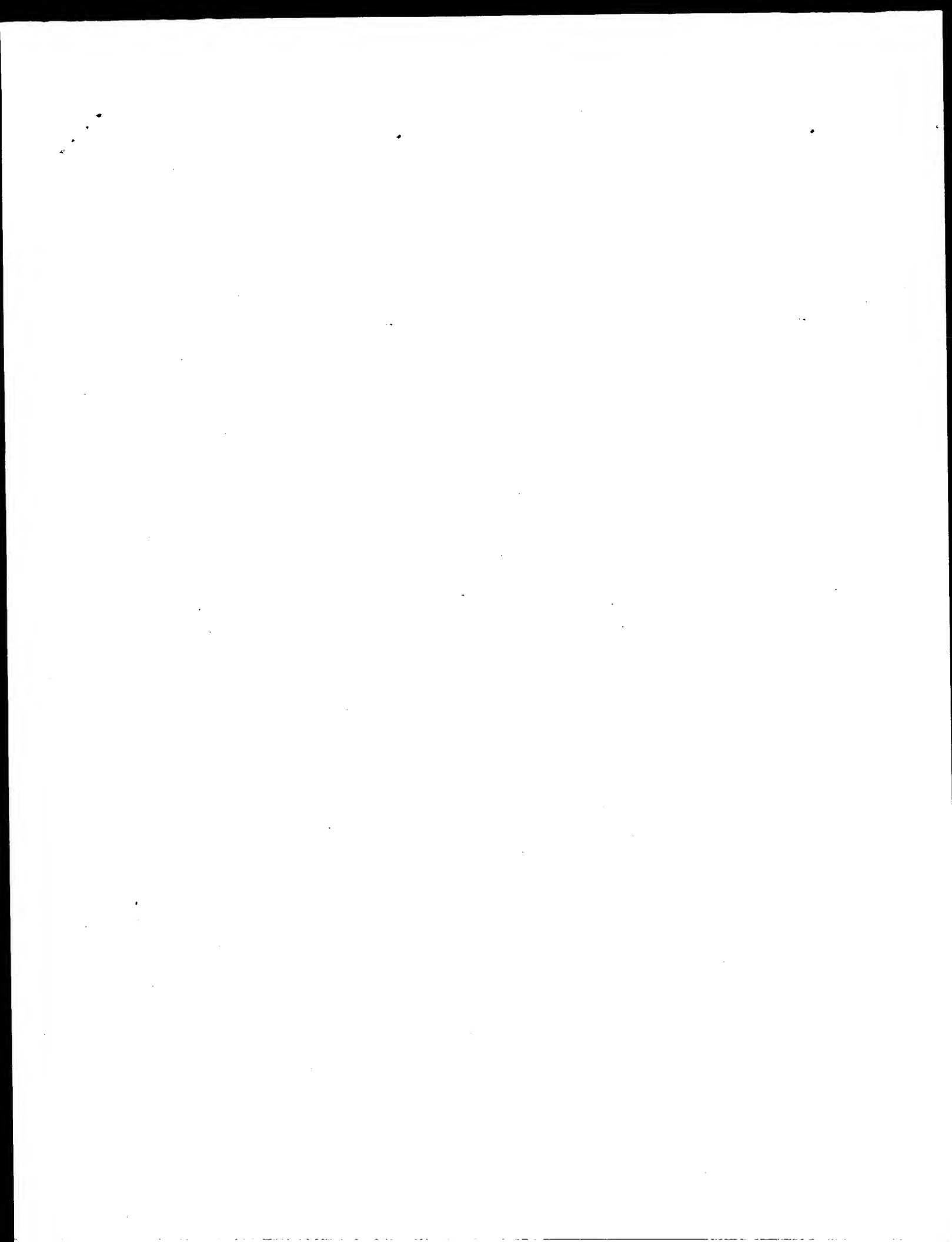
Qy 1 IFSGNNMIKMK 11
 Db 139 IFSGNNMIKMK 149

Sat May 10 08:44:08 2003

us-09-142-524d-29.rsp

Page 9

Search completed: April 20, 2003, 13:07:32
Job time : 5.92105 secs



GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 12.6711 Seconds
(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524D-29

Perfect score: 79

Sequence: 1 IFSGNNMIKIMPMY 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phase:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_unclassified:*
15: sp_rvitus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	374	10 Q8RUR1	Q8RUR1 cryptomeria
2	68	86.1	367	10 Q9LIT2	Q9LIT2 juniperus v
3	68	86.1	367	10 Q9LIT1	Q9LIT1 juniperus v
4	68	86.1	367	10 Q9M4S6	Q9M4S6 cupressus s
5	68	86.1	367	10 Q9M4S4	Q9M4S4 cupressus s
6	68	86.1	367	10 Q9M4S3	Q9M4S3 cupressus s
7	68	86.1	367	10 Q9M4S2	Q9M4S2 cupressus s
8	68	86.1	367	10 Q9M4S1	Q9M4S1 juniperus o
9	64	81.0	367	10 Q9M4S5	Q9M4S5 cupressus s
10	64	81.0	367	10 Q9M4S6	Q9M4S6 cupressus s
11	42	53.2	272	8 Q9B295	Q9B295 ceratopsolen
12	42	53.2	312	11 Q8VG18	Q8VG18 mus musculu
13	42	53.2	521	8 Q9B975	Q9B975 ceratopsolen
14	42	53.2	543	12 Q83893	Q83893 ovine adeno
15	40	50.6	245	12 Q84447	Q84447 paramacium
16	40	50.6	262	8 Q21044	Q21044 dictyostell

17	40	50.6	341	16 Q9Z8Q8	Q9Z8Q8 chlamydia p
18	40	50.6	445	9 Q8SD05	Q8SD05 pseudomonas
19	40	50.6	578	16 Q9X9X5	Q9X9X5 staphylococ
20	40	50.6	888	5 Q9S080	Q9S080 drosophila
21	40	50.6	1069	5 Q9VF64	Q9VF64 drosophila
22	40	50.6	1084	5 Q8SWR8	Q8SWR8 drosophila
23	39	49.4	306	2 Q9L973	Q9L973 lactococcus
24	39	49.4	376	16 Q9CMT9	Q9CMT9 pasteurella
25	39	49.4	421	11 Q9QYR8	Q9QYR8 mus musculu
26	39	49.4	766	5 Q25823	Q25823 plasmodium
27	39	49.4	794	16 Q8XB12	Q8XB12 escherichia
28	39	49.4	1611	16 Q8RFV3	Q8RFV3 fusobacteri
29	38	48.1	272	8 Q9G7X0	Q9G7X0 meselatus f
30	38	48.1	283	17 Q97YE1	Q97YE1 sulfolobus
31	38	48.1	304	2 Q9EVT5	Q9EVT5 thermotoga
32	38	48.1	310	2 Q93L00	Q93L00 bacillus th
33	38	48.1	316	16 Q8RH2	Q8RH2 fusobacteri
34	38	48.1	326	11 Q8VFR1	Q8VFR1 mus musculu
35	38	48.1	347	8 Q8RV6	Q8RV6 guillardia
36	38	48.1	369	2 Q8RSW6	Q8RSW6 synchococc
37	38	48.1	371	10 Q9AF00	Q9AF00 arabidopsis
38	38	48.1	440	12 Q9EMV0	Q9EMV0 amesacta moo
39	38	48.1	442	2 Q9KX32	Q9KX32 streptococ
40	38	48.1	488	10 Q22962	Q22962 arabidopsis
41	38	48.1	501	16 Q8RB7	Q8RB7 thermoaer
42	38	48.1	757	5 Q21407	Q21407 caenorhabd
43	38	48.1	992	11 Q9WU50	Q9WU50 mus musculu
44	38	48.1	1007	5 Q9U0D1	Q9U0D1 aplysia cal
45	38	48.1	1344	11 Q35851	Q35851 mus musculu

ALIGNMENTS

RESULT 1

Q8RUR1

ID Q8RUR1 PRELIMINARY; PRT; 374 AA.

AC Q8RUR1; 01-JUN-2002 (TREMUR1.21, Created)

DT 01-JUN-2002 (TREMUR1.21, Last sequence update)

DT 01-JUN-2002 (TREMUR1.21, Last annotation update)

DE Cry j 1 precursor.

GN Cry j 1.1 OR Cry j 1.2.

OS Cryptomeria japonica (Japanese cedar).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.

OX NCBI_TaxID=3369;

RN [1]

RP SEQUENCE FROM N.A.

RC TRISUB=POLLEN;

RA Futamura N., Shinohara K.;

RT "Isolation and characterization of cDNAs encoding major allergen Cry j

1 from Cryptomeria japonica pollen."

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB081309; BAB86286.1; -

DR EMBL; AB081310; BAB86287.1; -

KM Signal.

FT SIGNAL 1 21 POTENTIAL. CRY J 1.

FT CHAIN 22 374

SQ SEQUENCE 374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;

Query Match 100.0%; Score 79; DB 10; Length 374;

Best Local Similarity 100.0%; Pred. No. 2.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFSGNNMIKIMPMY 15

Db 92 IFSGNNMIKIMPMY 106

RESULT 2

ID Q9LIT2 PRELIMINARY; PRT; 367 AA.

```

AC 09LTL2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-2.
OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
NCBI_TaxID=39584;
RN
RP SEQUENCE FROM N.A.
RA MEDLINE=2135424; PubMed=11422137;
RA Midoro-Horvut T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778(2001).
DR EMBL; AF151427; AAF80164.1; -
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF69586 CRC64;

Query Match 86.1%; Score 68; DB 10; Length 367;
Best Local Similarity 86.7%; Pred. No. 0.00028;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
DB 92 IFSGNNNIKLKMPY 106

RESULT 3
09LTL1 PRELIMINARY; PRT; 367 AA.
ID 09LTL1;
AC 09LTL1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-1.
OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
NCBI_TaxID=39584;
RN
RP SEQUENCE FROM N.A.
RA MEDLINE=2135424; PubMed=11422137;
RA Midoro-Horvut T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778(2001).
DR EMBL; AF151429; AAF80166.1; -
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;

Query Match 86.1%; Score 68; DB 10; Length 367;
Best Local Similarity 86.7%; Pred. No. 0.00028;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
DB 92 IFSGNNNIKLKMPY 106

RESULT 4
09M4S6 PRELIMINARY; PRT; 367 AA.
ID 09M4S6;
AC 09M4S6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 Pollen allergen.

```

```

GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
NCBI_TaxID=13469;
RN
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257491; AAF72625.1; -
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39835 MW; BA59CC0108C25A3 CRC64;

Query Match 86.1%; Score 68; DB 10; Length 367;
Best Local Similarity 86.7%; Pred. No. 0.00028;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
DB 92 IFSGNNNIKLKMPY 106

RESULT 5
09M4S4 PRELIMINARY; PRT; 367 AA.
ID 09M4S4;
AC 09M4S4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
NCBI_TaxID=13469;
RN
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257493; AAF72627.1; -
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 86.1%; Score 68; DB 10; Length 367;
Best Local Similarity 86.7%; Pred. No. 0.00028;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
DB 92 IFSGNNNIKLKMPY 106

RESULT 6
09M4S3 PRELIMINARY; PRT; 367 AA.
ID 09M4S3;
AC 09M4S3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
NCBI_TaxID=13469;

```


RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257494; AAF72628.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase1.1.
 DR PRINTS; PRO0807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39832 MW; B5DFBF5A61C07A53 CRC64;

Query Match 86.1%; Score 68; DB 10; Length 367;
 Best Local Similarity 86.7%; Pred. No. 0.00028;
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
 DB 92 IFSGNNNIKLKMPY 106

RESULT 7
 ID Q9M4S2 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSL.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257495; AAF72629.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase1.1.
 DR PRINTS; PRO0807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39819 MW; AE7E055A61C07A53 CRC64;

Query Match 86.1%; Score 68; DB 10; Length 367;
 Best Local Similarity 86.7%; Pred. No. 0.00028;
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
 DB 92 IFSGNNNIKLKMPY 106

RESULT 8
 ID Q93X51 PRELIMINARY; PRT; 367 AA.
 AC Q93X51;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Putative allergen Jun o 1.
 GN JUN O 1.
 OS Juniperus oxycedrus (Prickly Juniper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxID=99008;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=POLLEN;
 RC TISSUE=POLLEN;
 FT TISSUE=POLLEN;
 RT "Cloning of Juniperus oxycedrus major allergen.";

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ293767; CAC48400.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase1.1.
 SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBFDB19D7 CRC64;

Query Match 86.1%; Score 68; DB 10; Length 367;
 Best Local Similarity 86.7%; Pred. No. 0.00028;
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
 DB 92 IFSGNNNIKLKMPY 106

RESULT 9
 ID Q9M4S5 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSL.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257492; AAF72626.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase1.1.
 DR PRINTS; PRO0807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0E3283B741 CRC64;

Query Match 81.0%; Score 64; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.0015;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
 DB 92 IFSGNNNIKLKMPY 106

RESULT 10
 ID Q93XL6 PRELIMINARY; PRT; 367 AA.
 AC Q93XL6;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative allergen Cup a 1 precursor.
 GN CUP A 1.
 OS Cupressus arizonica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=49011;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=POLLEN;
 RC TISSUE=POLLEN;
 FT TISSUE=POLLEN;
 RT "Cloning of Cupressus arizonica major allergen."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ278498; CAC37790.2; -
 DR Signal.
 KW Signal.
 FT CHAIN
 SQ SEQUENCE 367 AA; 39809 MW; AEP97260423A9F28 CRC64;

Query Match 81.0%; Score 64; DB 10; Length 367;
 Best Local Similarity 80.0%; Pred. No. 0.0015;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNIKLKMPLY 15
 Db 92 IFSGNNIKLKMPLY 106

RESULT 11

Q9B295 PRELIMINARY; PRT; 272 AA.
 ID Q9B295;
 AC Q9B295;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1) (Fragment).
 GN Cyt.
 OS Ceratopsolen bisulcatus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 OC Chalcidoidea; Agaonidae; Agaoninae; Ceratopsolen.
 OX NCBI_TaxID=130004;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C.BISULC;
 RX MEDLINE=21217342; PubMed=11321056;
 RA Machado C.A., Jouselin R., Kjellberg F., Compton S.G., Herre E.A.;
 RT "Phylogenetic relationships, historical biogeography, and character
 evolution of fig pollinating wasps."
 RL Proc. R. Soc. Lond., B, Biol. Sci. 268:685-694 (2001).
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
 CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 C + 2 H(2)O.
 CC -1- COFACTOR: HEMES A, A3, AND COPPER B (BY SIMILARITY).
 CC -1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC -1- INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AY014986; AAK13227.1; -.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.
 KW Copper; Heme; Inner membrane; Mitochondrion; Oxidoreductase;
 KM Respiratory chain; Transmembrane.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 272 AA; 30878 MW; E78E6C72148A1867 CRC64;

Query Match 53.2%; Score 42; DB 8; Length 272;
 Best Local Similarity 64.3%; Pred. No. 13;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IFSGNNIKLKMPLY 14
 Db 242 IFSGNNIKLKMPLY 255

RESULT 12

Q8VG18 PRELIMINARY; PRT; 312 AA.
 ID Q8VG18;
 AC Q8VG18;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Olfactory receptor MOR112-2.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang X., Firestein S.J.;
 RT "The olfactory receptor gene superfamily of the mouse."
 RL Nat. Neurosci. 0:0-0 (2002).

RESULT 13

Q9B975 PRELIMINARY; PRT; 521 AA.
 ID Q9B975;
 AC Q9B975;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1) (Fragment).
 OS Ceratopsolen bisulcatus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 OC Chalcidoidea; Agaonidae; Agaoninae; Ceratopsolen.
 OX NCBI_TaxID=130004;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B170;
 RX Weiblen G.D.;
 RA "Phylogenetic analyses of dioecious fig pollinators based on
 mitochondrial DNA sequences and morphology."
 RT Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
 CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 AND COPPER B (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME
 C + 2 H(2)O.
 CC -1- COFACTOR: HEMES A, A3, AND COPPER B (BY SIMILARITY).
 CC -1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC -1- INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 DR EMBL; AF200375; AAK0062.1; -.
 DR HSRP; P18401; 1PRT.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 DR PROSITE; PS00077; COX1; 1.
 KW Copper; Heme; Inner membrane; Mitochondrion; Oxidoreductase;

Query Match 53.2%; Score 42; DB 11; Length 312;
 Best Local Similarity 47.6%; Pred. No. 15;
 Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

QY 3 SGNN-----NIKLMKMPY 15
 Db 38 SGNNITTTTLNLSNVAHKLTPMY 58

KM Respiratory chain; Transmembrane.
 PT NON-TER 1
 SQ SEQUENCE 521 AA; 58771 MW; 164C7EF2C4B66C5B CRC64;

Query Match 53.2%; Score 42; DB 8; Length 521;
 Best Local Similarity 64.3%; Pred. No. 24;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IFSGNNMKLKM 14
 DB 464 IFSGNNMKLKM 477

RESULT 14

ID 083893 PRELIMINARY; PRT; 543 AA.

AD 083893;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Fiber.

OS Ovine adenovirus type 7.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Atadenovirus.
 OX NCBI_TaxID=114430;

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=OAV287;
 RX MEDLINE=95297141; PubMed=7778275;

RA Virati S., Boyle D., Kocherhans R., Both G.W.;

RT "Sequence of ovine adenovirus homologs for 100K hexon assembly, 33K, pVIII, and fiber genes: early region E3 is not in the expected location";
 RL Virology 209:400-408(1995).

RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=OAV287;
 RX MEDLINE=96240641; PubMed=8659111;

RA Virati S., Brookes D.E., Strike P., Khatri A., Boyle D.B., Both G.W.;

RT "Unique genome arrangement of an ovine adenovirus: identification of new proteins and proteinase cleavage sites";
 RL Virology 220:186-199(1996).

RN [3]
 RP SEQUENCE FROM N.A.

RC STRAIN=OAV287;
 RX MEDLINE=97271311; PubMed=9126262;

RA Xu Z., Hyatt A., Boyle D.B., Both G.W.;

RT "Construction of ovine adenovirus recombinants by gene insertion or deletion of related terminal region sequences";
 RL Virology 230:62-71(1997).

RN [4]
 RP SEQUENCE FROM N.A.

RC STRAIN=OAV287;
 RX MEDLINE=98277079; PubMed=9614874;

RA Khatri A., Both G.W.;

RT "Identification of transcripts and promoter regions of ovine adenovirus OAV287";
 RL Virology 245:128-141(1998).

DR EMBL; U40839; AAA73919.1;
 DR InterPro; IPR000939; Adeno_fiber2.
 DR Pfam; PF00608; adeno_fiber2; 10.

SQ SEQUENCE 543 AA; 58217 MW; 20FA2F5CCTA8C915 CRC64;

Query Match 53.2%; Score 42; DB 12; Length 543;
 Best Local Similarity 72.7%; Pred. No. 25;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GNNMKLKM 14
 DB 178 GNNMKLKM 188

RESULT 15
 084447

ID 084447 PRELIMINARY; PRT; 245 AA.

AC 084447;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE A127R protein.

RN A127R.

OS Paramyxium bursaria chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
 OX NCBI_TaxID=10506;

RN [1]
 RP SEQUENCE FROM N.A.

RC MEDLINE=95407089; PubMed=7676624;
 RA Li Y., Lu Z., Burbank D.E., Kutish G.F., Rock D.L., Van Etten J.L.;

RT "Analysis of 43 kb of the Chlorella virus PBCV-1 330-kb genome: map positions 45 to 88.";
 RL Virology 212:134-150(1995).

RN [2]
 RP SEQUENCE FROM N.A.

RC MEDLINE=20013326; PubMed=10544099;
 RA Kaiser A., Volmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.;

RT "Chlorella virus PBCV-1 encodes a functional homospertidine synthase.";
 RL Virology 263:254-262(1999).

RN [3]
 RP SEQUENCE FROM N.A.

RC MEDLINE=20478054; PubMed=11021991;
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;

RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus PBCV-1.";
 RL Virology 276:27-36(2000).

RN [4]
 RP SEQUENCE FROM N.A.

RC Van Etten J.L.;

RT Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.

RN [5]
 RP SEQUENCE FROM N.A.

RC Van Etten J.L.;

RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RN [6]
 RP SEQUENCE FROM N.A.

RC Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

RN [7]
 RP SEQUENCE FROM N.A.

RC Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

RN [8]
 RP SEQUENCE FROM N.A.

RC Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [9]
 RP SEQUENCE FROM N.A.

RC Gurnon J.R., Graves M.V., Van Etten J.L.;

RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; U42580; AAC96495.1;
 DR InterPro; IPR000939; Adeno_fiber2.
 DR Pfam; PF00608; adeno_fiber2; 10.

SQ SEQUENCE 245 AA; 27126 MW; C8DC76A169CDE6 CRC64;

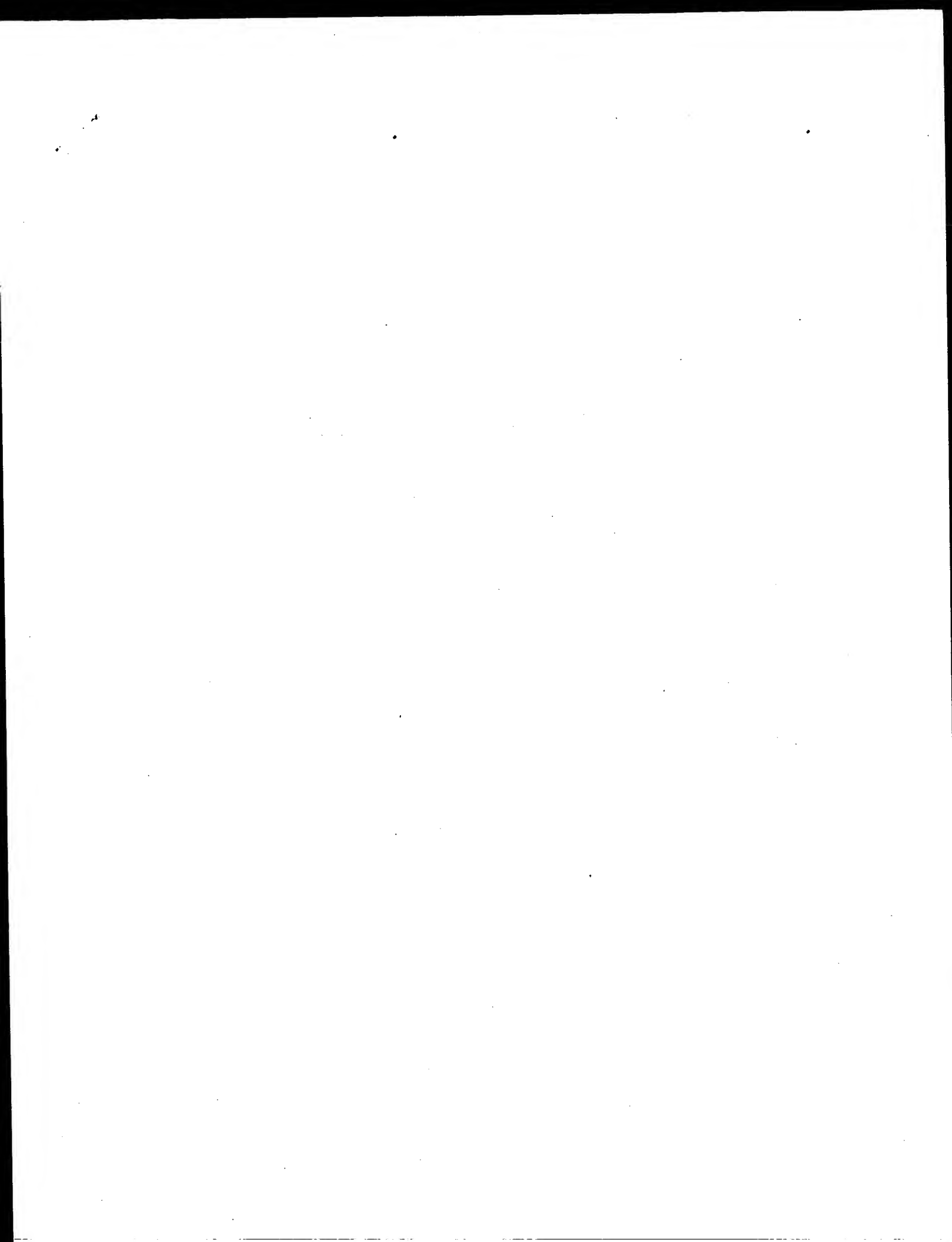
Query Match 50.6%; Score 40; DB 12; Length 245;
 Best Local Similarity 72.7%; Pred. No. 27;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FSGNNMKLKM 12
 DB 62 FSGNNMKLKM 72

Search completed: April 20, 2003, 13:13:01

Job time : 14.6711 secs



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524d-36

Sequence: 1 PCVPIKRVSNVTHIG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	374	1	SBP_CRYUA
2	58	72.5	375	1	MPAI_CHAAB
3	57	71.2	346	1	MPAI_CUPAR
4	55	68.8	367	1	MPAI_TUNAS
5	55	68.8	404	1	9612_LYCS
6	42	52.5	290	1	KEPR_NITVU
7	41	51.2	230	1	SFSA_PYRVU
8	39	48.8	439	1	LIP1_DROME
9	39	48.8	720	1	YMS2_YEAST
10	38	47.5	158	1	PRPB_ECOLI
11	38	47.5	214	1	Y023_CABEL
12	38	47.5	387	1	REL_MOUSE
13	38	47.5	1328	1	POLX_TOBAC
14	37	46.2	146	1	SP22_BACSH
15	37	46.2	290	1	ATNC_HUMAN
16	37	46.2	290	1	ATNC_MOUSE
17	37	46.2	290	1	ATNC_MOUSE
18	37	46.2	315	1	TRXB_MYGE
19	37	46.2	326	1	PEL_EMENT
20	37	46.2	449	1	PEB5_LYCS
21	37	46.2	452	1	IEB3_VZVD
22	37	46.2	722	1	VGLH_GPCMV
23	37	46.2	985	1	NAH1_YEAST
24	37	46.2	4128	1	PRKD_MOUSE
25	36	45.0	124	1	FOLB_CHLTR
26	36	45.0	434	1	HIS8_CAMJE
27	36	45.0	364	1	PEL_DILLO
28	36	45.0	488	1	EXON_HSV6Z
29	36	45.0	551	1	TCPE_YEAST
30	36	45.0	601	1	YEO6_SCHPO
31	36	45.0	688	1	NRPI_CHICK
32	36	45.0	914	1	KRPI_YEAST
33	36	45.0	917	1	CTNA_DROME

ALIGNMENTS

RESULT 1	ID	SBP_CRYUA	STANDARD	PRT	374 AA.
AC	P18632				
DT	01-NOV-1990	(Rel. 16, Created)			
DT	01-NOV-1995	(Rel. 32, Last sequence update)			
DT	30-MAY-2000	(Rel. 39, Last annotation update)			
DE	Sugi basic protein precursor (SBP) (Major allergen Cry j 1) (Cry j 1).				
OS	Cryptomeria japonica (Japanese cedar).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Coniferales; Cupressales; Cryptomeria.				
OX	NCBI_TaxID=3369;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE=Pollen;				
RX	MEDLINE=94183234; PubMed=8135802;				
RA	Sone T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K.,				
RA	Kino K.,				
RT	Cloning and sequencing of cDNA coding for Cry j I, a major allergen				
RT	of Japanese cedar pollen."				
RL	Biochem. Biophys. Res. Commun. 199;619-625(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Pollen;				
RA	Namba M., Kurose M., Torigoe K., Fukuda S., Kurimoto M.;				
RL	Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE OF 22-41.				
RC	TISSUE=Pollen;				
RX	MEDLINE=89031257; PubMed=3181436;				
RA	Tanai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inoue S.,				
RA	Matunasi T.;				
RT	N-terminal amino acid sequence of a major allergen of Japanese cedar				
RT	pollen (Cry j I)."				
RL	FEBS Lett. 239;332-332(1988).				
RN	[4]				
RP	CARBOHYDRATES.				
RC	TISSUE=Pollen;				
RX	MEDLINE=95003748; PubMed=7920021;				
RA	Hijikata A., Matsumoto I., Kojima K., Ogawa H.;				
RT	Antigenicity of the oligosaccharide moiety of the Japanese cedar				
RT	(Cryptomeria japonica) pollen allergen, Cry j I."				
RL	Int. Arch. Allergy Immunol. 105;196-202(1994).				
RN	[5]				
RP	STRUCTURE OF CARBOHYDRATES.				
RC	TISSUE=Pollen;				
RX	MEDLINE=95332249; PubMed=7608114;				
RA	Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Usui M.,				
RA	Fukuda S., Hanzawa H., Hanyama H., Kurimoto M.;				
RT	Carbohydrate structures of the glycoprotein allergen Cry j I from				
RT	Japanese cedar (Cryptomeria japonica) pollen."				
RL	J. Biochem. 117;289-295(1995).				
CC	-1- PTM: CONTAINS FUCOSE/XLOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.				
CC	-1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR				
CC	POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN.				
CC	-1- MISCELLANEOUS: THE SEQUENCE OF CRY j I FORM A IS SHOWN HERE. FORM				

CC B DIFFERS IN SIX POSITIONS.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D26544; BAA05542.1; -
 CC EMBL; D26544; BAA05543.1; -
 CC EMBL; D34639; BAA07020.1; -
 CC PIR; A44773; A44773.
 CC DR GLYCOSITEDB; P18632; -
 CC InterPro; IPR002022; Amb_allergen.
 CC Pfam; PF00544; pec_lyase; 1.
 CC DR PRINTS; PR00807; AMBALLERGEN.
 CC Allergen; Glycoprotein; Multigene family; signal.
 CC KW SIGNAL 1 21
 CC FT CHAIN 1 21
 CC FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (COMPLEX).
 CC FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (COMPLEX).
 CC FT CARBOHYD 354 354 /FTID-CAR 000135.
 CC FT CARBOHYD 354 354 /FTID-CAR 000136.
 CC FT VARIANT 12 12 L -> F (IN CRY J 1-B).
 CC FT VARIANT 143 143 H -> Y (IN CRY J 1-B).
 CC FT VARIANT 202 202 S -> T (IN CRY J 1-B).
 CC FT VARIANT 221 221 L -> S (IN CRY J 1-B).
 CC FT VARIANT 358 358 Q -> H (IN CRY J 1-B).
 CC FT VARIANT 361 361 K -> Q (IN CRY J 1-B).
 CC FT VARIANT 361 361
 CC SEQUENCE 374 AA; 40645 MW; 74AB25950248F56F CRC64;
 CC
 CC Query Match 100.0%; Score 80; DB 1; Length 374;
 CC Best Local Similarity 100.0%; Pred. No. 6.6e-07;
 CC Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 PCVFIRKSVNVIHG 15
 CC DB 127 PCVFIRKSVNVIHG 141
 CC
 CC RESULT 2
 CC MPAL CHAOB STANDARD; PRT; 375 AA.
 CC ID MPAL CHAOB
 CC AC 096385;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Major pollen allergen Cha o 1 precursor.
 CC OS Chamaecyparis obtusa (Japanese cypress).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
 CC Chamaecyparis.
 CC NCBI_TaxID=13415;
 CC [1]
 CC RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC RP TISSUE=Pollen;
 CC RA MEDLINE=96255194; PubMed=8676896;
 CC RA Suzuki M., Komiyama N., Itoh H., Sone T., Kuno K., Takagi I.,
 CC Ohta N.;
 CC RT "Purification, characterization and molecular cloning of Cha o 1, a
 CC major allergen of Chamaecyparis obtusa (Japanese cypress) pollen.";
 CC RT Mol. Immunol. 33:451-460(1996).
 CC RT
 CC CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D45404; BAA08246.1; -
 CC InterPro; IPR002022; Amb_allergen.
 CC Pfam; PF00544; pec_lyase; 1.
 CC DR PRINTS; PR00807; AMBALLERGEN.
 CC Allergen; Glycoprotein; signal.
 CC KW SIGNAL 1 21
 CC FT CHAIN 1 21
 CC FT CARBOHYD 22 375 MAJOR POLLEN ALLERGEN CHA O 1.
 CC FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 352 352
 CC SEQUENCE 375 AA; 40258 MW; 81CD91DF7066DBBF CRC64;
 CC
 CC Query Match 72.5%; Score 58; DB 1; Length 375;
 CC Best Local Similarity 60.0%; Pred. No. 0.006; 1; Indels 0; Gaps 0;
 CC Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 PCVFIRKSVNVIHG 15
 CC DB 127 PCVFIRKSVNVIHG 141
 CC
 CC RESULT 3
 CC MPAL CUPAR STANDARD; PRT; 346 AA.
 CC ID MPAL CUPAR
 CC AC 095C9;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Major pollen allergen Cup a 1.
 CC OS Cupressus arizonica.
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 CC NCBI_TaxID=99011;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP MEDLINE=20571526; PubMed=11122214;
 CC RA Aceituno E., Del Pozo V., Minquez A., Arrieta I., Cortegano I.,
 CC RA Cardaba B., Gallardo S., Rojo M., Palomero P., Lahoz C.;
 CC RT "Molecular cloning of major allergen from Cupressus arizonica pollen:
 CC Cup a 1.";
 CC RT Clin. Exp. Allergy 30:1750-1758(2000).
 CC RT
 CC CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ243570; CAB62551.1; -
 CC InterPro; IPR002022; Amb_allergen.
 CC Pfam; PF00544; pec_lyase; 1.
 CC DR PRINTS; PR00807; AMBALLERGEN.
 CC Allergen; Glycoprotein.
 CC KW CARBOHYD 127 127
 CC FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 272 272
 CC SEQUENCE 346 AA; 37589 MW; F1281DCDA1D5DFD0 CRC64;
 CC
 CC Query Match 71.2%; Score 57; DB 1; Length 346;
 CC Best Local Similarity 53.3%; Pred. No. 0.0084; 1; Indels 0; Gaps 0;
 CC Matches 8; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 1 PCVFKRVSNTIIG 15
 DB 106 PCLFMRKASHVILHG 120

RESULT 4

MPAL JUNAS STANDARD; PRT; 367 AA.
 ID MPAL JUNAS
 AC P81294; 09ZNU7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 OS Major pollen allergen Jun a 1 precursor.
 DE Juniperus ashei (Ozark white cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Cupressaceae; Juniperus.
 NCBI_TaxID=11101;

SEQUENCE FROM N.A., AND SEQUENCE OF 31-38; 42-50; 58-80; 88-94;
 RP P81294; 134-140; 160-164; 256-263 AND 322-325.
 RC TISSUE=Pollen;
 RA MEDLINE=99414163; PubMed=10482836;
 RA Midozo-Horintu T.M., Goldblum R.M., Kurosky A., Wood T.G.,
 RA Brooks E.G.;
 RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 RT allergen, Jun a 1."
 RL J. Allergy Clin. Immunol. 104:613-617(1999).
 RN [2]

SEQUENCE OF 22-50.
 RP TISSUE=Pollen;
 RA MEDLINE=99414162; PubMed=10482835;
 RA Midozo-Horintu T., Goldblum R.M., Kurosky A., Goetz D.W.,
 RA Brooks E.G.;
 RT "Isolation and characterization of the mountain cedar (Juniperus
 RT ashei) pollen major allergen, Jun a 1."
 RL J. Allergy Clin. Immunol. 104:608-612(1999).
 CC -1- DISEASE: THIS PROTEIN CAUSES SEVERAL SEASONAL ALLERGIC RHINITIS
 CC IN NORTH AMERICA.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AF106663; AAD03609.1; -
 DR EMBL; AF106662; AAD03608.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase1.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KM Allergen; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 367 MAJOR POLLEN ALLERGEN JUN A 1.
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 367 AA; 39824 MW; FC9B81E675662E49 CRC64;

Query Match 68.8%; Score 55; DB 1; Length 367;
 Best Local Similarity 57.1%; Pred. No. 0.02;
 Matches 8; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFKRVSNTIIG 14
 DB 127 PCLFMRKASHVILHG 140

RESULT 5
 9612_LYCES

ID 9612 LYCES STANDARD; PRT; 404 AA.
 AC P24396;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-1999 (Rel. 38, Last annotation update)
 DE Style development-specific protein 9612 precursor.
 GN 9612.

OS Lycopersicon esculentum (Tomato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4081;
 RN [1]

SEQUENCE FROM N.A.
 RP STRAIN=cv. VF36; TISSUE=Pistil;
 RC MEDLINE=91117185; PubMed=2277637;
 RA Budellier K.A., Smith A.G., Gasser C.S.;
 RT "Regulation of a stylar transmitting tissue-specific gene in
 RT wild-type and transgenic tomato and tobacco."
 RL Mol. Gen. Genet. 224:183-192(1990).
 CC -1- FUNCTION: MAY HAVE A ROLE IN THE DEVELOPMENT OF THE TRANSMITTING
 CC TISSUE OF THE STYLE AND/OR IN THE EVENTS RELATED TO POLLINATION
 CC SUCH AS SOME ASPECT IN THE FACILITATION OF COMPATIBLE POLLEN TUBE
 CC GROWTH.

CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN THE PISTIL WHERE IT IS
 CC FOUND IN THE OUTER FIVE LAYERS OF THE STRANDS OF TRANSMITTING
 CC TISSUE WITHIN THE UPPER TWO-THIRDS OF THE STYLE. FOUND AT MUCH
 CC LOWER LEVELS IN THE ANTHERS AND VEGETATIVE ORGANS.
 CC -1- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND DURING ANTHERESIS.
 CC -1- SIMILARITY: 54% IDENTICAL TO TOMATO PROTEINS P59 (AC P15722)
 CC AND P56 (AC P15721).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; X55193; CAA8979.1; -
 DR PIR; S12209; S12209.
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase1.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KM Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 404
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 404 AA; 44298 MW; B2CED6B128D8675 CRC64;

Query Match 68.8%; Score 55; DB 1; Length 404;
 Best Local Similarity 60.0%; Pred. No. 0.022;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 PCVFKRVSNTIIG 15
 DB 158 PCLTHHTSNTIIG 172

RESULT 6
 KPRP NITVU STANDARD; PRT; 290 AA.
 ID KPRP NITVU
 AC P37100;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphoribulokinase (EC 2.7.1.19) (Phosphopentokinase) (PRK).
 GN CHRP
 OS Nitrobacter vulgaris.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

CC Bradyrhizobium group; Nitrobacteraceae; Nitrobacter.
 OX NCBI_TaxID=29421;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=T3;
 RA Strecker M., Sickinger E., English R.S., Shively J.M., Bock E.;
 RL Submitted (XX-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + D-ribulose 5-phosphate = ADP + D-ribulose 1,5-bisphosphate.
 CC -1- PATHWAY: Calvin cycle.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHORIBULOKINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L22884; AAA25506.1; -
 CC HSSP; P12033; 1A7G.
 DR InterPro; IPR001324; PRK.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00478; PHRIBKINASE.
 DR PROSITE; PS00567; PHOSPHORIBULOKINASE; 1.
 DR Trasnferase; Kinase; Calvin cycle; ATP-binding; Photosynthesis.
 KW NP BIND 12
 FT ATP (POTENTIAL).
 SQ SEQUENCE 290 AA; 32908 MW; 9E29722319A2B41C CRC64;
 Query Match 52.5%; Score 42; DB 1; Length 290;
 Best Local Similarity 40.0%; Pred. No. 3.6;
 Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 PCVFIKRVNVIHG 15
 Db 251 PSSFMRANSIVIHG 265
 RESULT 7
 ID SFSA_PYRFU STANDARD; PRT; 230 AA.
 AC Q8U1K8;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sugar fermentation stimulation protein homolog.
 GN SFSA OR PF1198.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Yc1 / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RL -1- SIMILARITY: BELONGS TO THE SFSA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE010228; AAL81322.1; -
 DR Complete proteome.
 KW SEQUENCE 230 AA; 26114 MW; 41B3D94D68C7A717 CRC64;
 SQ SEQUENCE 230 AA; 26114 MW; 41B3D94D68C7A717 CRC64;
 Query Match. 51.2%; Score 41; DB 1; Length 230;

Best Local Similarity 41.7%; Pred. No. 4.3;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 PCVFIKRVNVI 12
 Db 11 PCVFIKRVNRFV 22
 RESULT 8
 ID LIPI_DROME STANDARD; PRT; 439 AA.
 AC Q46107; Q9VNR6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lipase 1 precursor (EC 3.1.1.-).
 GN LIPI OR CG7279.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RN SEQUENCE OF 7-433 FROM N.A.
 RC STRAIN=Canton-S;
 RX MEDLINE=98227315; PubMed=9566193;
 RA Piatillo D., Marzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;
 RT "The Drosophila melanogaster lipase homologs: a gene family with
 RT tissue and developmental specific expression.";
 RL J. Mol. Biol. 276:877-885(1998).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoeklin R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle G., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballwey R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwan C.,
 RA Jaislin M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Paclab J.M.,
 RA Palazzolo M., Plattman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach U.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: COULD BE A DIGESTIVE ENZYME.

CC TISSUE SPECIFICITY: OVARIES.
 CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
 CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
 CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Y14366; CA474736.1; -
 CC EMBL; AB003629; AAF52994.1; -
 CC FlyBase; FBgn0023496; Lip1.
 CC InterPro; IPR000073; Abhydrolase.
 CC InterPro; IPR000374; Lipase.
 CC Pfam; PF00561; abhydrolase_1.
 CC DR PROSITE; PS00120; LIPASE_SER; FALSE NEG.
 CC Hydrolyase; Lipid degradation; Signal; Glycoprotein.
 CC KW SIGNAL 1 24
 CC FT CHAIN 25 439
 CC FT DOMAIN 30 44
 CC FT ACT_SITE 197 197
 CC FT ACT_SITE 393 393
 CC FT CARBOHYD 124 124
 CC FT CARBOHYD 151 151
 CC FT CARBOHYD 346 346
 CC FT CARBOHYD 379 379
 CC FT CARBOHYD 426 426
 CC FT CONFLICT 10 10
 CC FT CONFLICT 213 213
 CC FT CONFLICT 412 412
 CC FT SEQUENCE 439 AA; 50660 MW; 9E32E20BEAE93E3F CRC64;
 QY 2 CVFKRVSNTVING 15 48.8%; Score 39; DB 1; Length 439;
 Db 16 CVFISRIQGLIGG 29 50.0%; Pred. No. 18;
 Matches 7; Conservativity 2; Mismatches 5; Indels 0; Gaps 0;
 RESULT 9
 ID YMS2_YEAST STANDARD; PRT; 720 AA.
 AC Q04332;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 82.1 kDa protein in SGI-MRPL24 intergenic region.
 GN YMR192W OR YMR646.04.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Pearson D., Bowman S., Bartell B.G., Rajandream M.A.;
 RA Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO YEAST YPL249C AND S.POMBE SPAC66FL.09.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; Z47815; CA87813.1; -
 CC SGD; S0004804; YMR192W.
 CC InterPro; IPR00195; RabGAP_TBC.
 CC Pfam; PF00566; TBC; 1.
 CC SMART; SM00164; TBC; 1.
 CC Hypothetical protein.
 CC SEQUENCE 720 AA; 82131 MW; E1423DBA4F5F7267 CRC64;
 QY 1 PCVFKRVSNTV 12 48.8%; Score 39; DB 1; Length 720;
 Db 348 PSEYKRIISNVL 359 50.0%; Pred. No. 30;
 Matches 6; Conservativity 4; Mismatches 2; Indels 0; Gaps 0;
 RESULT 10
 ID PTPB_ECOLI STANDARD; PRT; 158 AA.
 AC P42909;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE PTS system, N-acetylglucosamine-specific IIB component 1 (E1B-AGA)
 DE (N-acetylglucosamine-permease IIB component 1) (Phosphotransferase
 DE enzyme II B component 1) (EC 2.7.1.69).
 GN AGAB OR B3138.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C;
 RC MEDLINE=20392444; PubMed=10931310;
 RA Brinkkoetter A., Kloess H., Alpert C.-A., Lengeler J.W.;
 RA "Pathways for the utilization of N-acetyl-galactosamine and
 RT galactosamine in Escherichia coli."; Mol. Microbiol. 37:125-135(2000).
 RL [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RC MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).
 RL [3]
 RN [3]
 RP DISCUSSION OF SEQUENCE.
 RC MEDLINE=97086503; PubMed=8932697;
 RA Reizer J., Rameisler T.M., Reizer A., Chardit A., Saler M.H. Jr.;
 RT "Novel phosphotransferase genes revealed by bacterial genome sequencing: a gene cluster encoding a putative N-acetylglucosamine
 RT metabolic pathway in Escherichia coli."; Microbiology 142:231-250(1996).
 RL [4]
 RN [4]
 RP FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
 CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE-DEPENDENT
 CC TRANSPORT SYSTEM. THE IICD DOMAINS CONTAIN THE SUGAR BINDING SITE
 CC AND THE TRANSMEMBRANE CHANNEL, THE IIA DOMAIN CONTAINS THE PRIMARY
 CC PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIB TRANSFERS ITS
 CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
 CC THE SUGAR.
 CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
 CC histidine + sugar phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: CONTAINS 1 PTS E1B DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF228498; AAF61090.1; -
DR EMBL: U18997; AAA57941.1; -
DR EMBL: AF000395; AAC76172.1; -
DR Ecogen: EG12769; agab.
DR InterPro: IPR004720; PTSIB scrb.
DR TIGR: TIGR00854; pbs-scrbse; 1.
KM Phosphotransferase system; Sugar transport; Transferase;
KW Phosphorylation; Complete proteome.
FT MOD RES 17 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 158 AA; 17622 MW; 897BAAD5FED89159 CRC64;

Query Match 47.5%; Score 38; DB 1; Length 158;
Best Local Similarity 33.3%; Pred. No. 10;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 PCVFIKRVSNVING 15
DB 4 PNTILTRIDNRLVHG 18

RESULT 11

Y023_CABEL STANDARD; PRT; 214 AA.

AC 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 23.3 kDa protein ZK688.3 in chromosome III.

GN ZK688.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Pelodermidae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Bristol N2.

KX MEDLINE=94150718; PubMed=796398;

RA Wilson R., Atkinson R., Anderson K., Baynes C., Betks M.,

RA Craton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,

RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,

RA Johnston L., Jones M., Kersey J., Kistner J., Laister N.,

RA Lattelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rinken L., Roopra A., Saunders D., Shownkeen R.,

RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer B., Staden R.,

RA Sultson J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,

RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,

RA Wohldman P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RT elegans."

RT Nature 368:32-38(1994).

RT RT

CC -1- SIMILARITY: BELONGS TO THE FAH / HPEC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

SQ SEQUENCE 214 AA; 23335 MW; 4C0CCEB3065D007 CRC64;
Query Match 47.5%; Score 38; DB 1; Length 214;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 PCVFIKRVSNVING 15
DB 36 PMLFVKTNSFIVEG 50

RESULT 12

REL_MOUSE STANDARD; PRT; 587 AA.

AC P15307;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-Rel proto-oncogene protein (C-Rel protein).

GN REL.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALE/c; TISSUE=Spleen;

RX MEDLINE=89239482; PubMed=2654811;

RA Grumont R.J., Gerondakis S.;

RT "Structure of a mammalian c-rel protein deduced from the nucleotide

RT sequence of murine CDNA clones.";

RL Oncogene Res. 4:1-8(1989).

RN [2]

RP REVISIONS.

RX MEDLINE=90370373; PubMed=2204017;

RA Grumont R.J., Gerondakis S.;

RT "The murine c-rel proto-oncogene encodes two mRNAs the expression of

RT which is modulated by lymphoid stimuli.";

RL Oncogene Res. 5:245-254(1990).

CC -1- FUNCTION: PROTO-ONCOGENE THAT MAY PLAY A ROLE IN DIFFERENTIATION

CC AND LYMPHOPOIESIS.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: X15842; CAA33843.1; ALT_SEQ.

DR EMBL: X60271; CAA42817.1; -

DR PIR: A45505; A45505.

DR PIR: A60367; A60367.

DR HSSP: P25799; 1BFT.

DR TRANSFAC: T00169; -

DR MGD: MGI:97897; Rel.

DR InterPro: IPR002909; IPT TIG.

DR InterPro: IPR000451; NF_Rel_dor_fam.

DR Pfam: PF00554; RHD; 1.

DR Pfam: PF01833; TIG; 1.

DR PRINTS: PR00057; NFKBNSCPCT.

DR SMART: SMO0429; IPT; 1.

DR PROSITE: PS01204; REL_1; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

DR PROSITE: PS50254; REL_2; 1.

Query Match 47.5%; Score 38; DB 1; Length 587;
 Best Local Similarity 58.3%; Pred. No. 37;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 FIKRSNVIIHG 15
 DB 327 FIKKESNUFSHG 338

RESULT 13

POLX TOBAC STANDARD; PRT; 1328 AA.
 ID POLX TOBAC
 AC P10978;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Retrovirus-related Pol polyprotein from transposon TMT-1-94 [Contains: Protease (EC 3.4.23.-); Reverse transcriptase (EC 2.7.7.49); Endonuclease].
 DE Nicotiana tabacum (Common tobacco).
 OS Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; easterids I; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89097311; PubMed=2536143;
 RA Gradshteyn M.-A., Spielmann A., Caboche M.;
 RT "Tmt1, a mobile retroviral-like transposable element of tobacco RT isolated by plant cell genetics."
 RL Nature 337:376-380 (1989).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + [DNA] (N).
 CC -1- SIMILARITY: HIGH, WITH DROSOPHILA COPA ELEMENT.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A11.
 CC -1- SIMILARITY: CONTAINS 1 CCHC-TYPE ZINC FINGER.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X13777; CAA32025.1; -
 DR PIR: S04273; S04273.
 DR MEROPS: A11.002; -
 DR InterPro: IPR001995; Aspprotease_rtrv.
 DR InterPro: IPR001584; Rve.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam: PF00077; rvd.1.
 DR Pfam: PF00098; zf-CCHC; 1.
 DR Pfam: PF00665; tve.1.
 DR PRINTS: PR00939; C2HCZNFINGER.
 DR SMART: SM00343; Znf_CCHC; 1.
 DR PROSITE: PS50158; ZF_CCHC; 1.
 KW Hydroxylase; Aspartyl protease; RNA-directed DNA polymerase; Endonuclease; Transferase; Polyprotein; Transposable element;
 KW Zinc-finger.
 FT ZN FING 230 247 CCHC-TYPE.
 FT ACT SITE 297 297 PROTEASE (BY SIMILARITY).
 SQ SEQUENCE 1328 AA; 151076 MW; F27E76C504B19B1B CRC64;

Query Match 47.5%; Score 38; DB 1; Length 1328;
 Best Local Similarity 60.0%; Pred. No. 81;
 Matches 9; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 PCVFIKRSV--NVII 13
 DB 990 PCVFIKRSVNNPFI 1004

RESULT 14
 SP22_BACSH STANDARD; PRT; 146 AA.
 ID SP22_BACSH
 AC O32724;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-1998 (Rel. 36, Last annotation update)
 DE Anti-sigma F factor (Stage II sporulation protein AB).
 GN SpoIIAB.
 OS Bacillus sphaericus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=1421;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2362;
 RX MEDLINE=97409954; PubMed=9266669;
 RA Park S.G., Yuckin M.D.;
 RT "Sequencing and phylogenetic analysis of the spoIIA operon from diverse Bacillus and Paenibacillus species."
 RL Gene 194:25-33 (1997).
 CC -1- FUNCTION: BINDS TO SIGMA F AND BLOCKS ITS ABILITY TO FORM AN RNA POLYMERASE HOLOENZYME (E-SIGMA F). PHOSPHORYLATES SPOIIA ON A SERINE RESIDUE. THIS PHOSPHORYLATION MAY ENABLE SPOIIA TO ACT AS AN ANTI-ANTI-SIGMA FACTOR THAT COUNTERACTS SPOIIAB AND THUS RELEASES SIGMA F FROM INHIBITION (BY SIMILARITY).
 CC -1- SIMILARITY: STRONG, TO RBM.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L47359; AAB81190.1; -
 DR InterPro: IPR003594; Atphind_ATPase.
 DR InterPro: IPR004359; HIS_KIN_sig.
 DR Pfam: PF02518; HATPase_c1.
 DR SMART: SM00387; HATPase_c1.
 KW Sporulation; Transferase; Kinase.
 SQ SEQUENCE 146 AA; 16124 MW; FB56849D67C20CAC CRC64;

Query Match 46.2%; Score 37; DB 1; Length 146;
 Best Local Similarity 61.5%; Pred. No. 14;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 VFIRKSNVVIHG 15
 DB 43 VVSEAVSNVVIHG 55

RESULT 15
 ATNC_HUMAN STANDARD; PRT; 290 AA.
 ID ATNC_HUMAN
 AC P14415; O60444;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sodium/potassium-transporting ATPase beta-2 chain (Sodium/potassium-dependent ATPase beta-2 subunit).
 GN ATP1B2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89174720; PubMed=2538450;
 RA Martin-Vasallo P., Dackowski W., Emanuel J.R., Levenson R.;
 RT "Identification of a putative isoform of the Na,K-ATPase beta subunit. Primary structure and tissue-specific expression."
 RL J. Biol. Chem. 264:4613-4618 (1989).

RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=94137737; PubMed=8305453;
 RA Hernandez N., Martin-Vasallo P., Ghosh S., Ghosh P.K., Swaroop A.,
 RA Coca-Prados M.;
 RT "Nucleotide sequence of a cDNA for the beta 2 subunit isoform of
 Na,K(+)ATPase from human retina."
 RL Biochim. Biophys. Acta 1189:109-111(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retinal pigment epithelium;
 RX MEDLINE=97075936; PubMed=8918259;
 RA Ruiz A.C., Bhat S.P., Bok D.;
 RT "Expression and synthesis of the Na,K-ATPase beta 2 subunit in human
 retinal pigment epithelium."
 RL Gene 176:237-242(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98201617; PubMed=9524271;
 RA Avila J., Alvarez de la Rosa D., Gonzalez-Martinez L.M., Lecuona E.,
 RA Martin-Vasallo P.;
 RT "Structure and expression of the human Na,K-ATPase beta 2-subunit
 gene."
 RL Gene 208:221-227(1998).
 CC -1- FUNCTION: THIS IS THE NON-CATALYTIC COMPONENT OF THE ACTIVE
 CC ENZYME, WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE
 CC EXCHANGE OF NA AND K IONS ACROSS THE PLASMA MEMBRANE. THE EXACT
 CC FUNCTION OF THIS GLYCOPROTEIN IS NOT KNOWN.
 CC -1- SUBUNIT: COMPOSED OF THREE SUBUNITS: ALPHA (CATALYTIC), BETA
 CC AND GAMMA.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
 CC -1- MISCELLANEOUS: THE BETA SUBUNIT SEEMS TO BE ENCODED BY A MULTIGENE
 CC FAMILY. EACH DIFFERENT SUBUNIT MAY HAVE SPECIALIZED FUNCTIONS.
 CC -1- SIMILARITY: BELONGS TO THE NA+/K+ AND H+ ATPASES BETA CHAIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M8181; AA51805.1; -
 CC EMBL; U45945; AAC50873.1; -
 CC EMBL; AF007876; AAC39686.1; -
 CC Gene; HGNC:805; ATP1B2.
 DR MIM; 182331; -
 DR InterPro; IPR000402; Na/K-ATPase_beta.
 DR Pfam; PF00287; Na_K-ATPase; 1.
 DR TIGRfam; TIGR01107; Na_K-ATPase_beta; 1.
 DR PROSITE; PS00390; ATPASE_NA_K_BETA_1; 1.
 DR PROSITE; PS00391; ATPASE_NA_K_BETA_2; 1.
 KW Sodium/potassium transport; Transmembrane; Glycoprotein;
 KW Multigene family; Signal-anchor.
 FT DOMAIN 1 39
 FT TRANSMEM 40 67
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT P -> I (IN REF. 3 AND 4).
 FT M -> I (IN REF. 3 AND 4).
 FT DOMAIN 68 290
 FT DISULFID 129 150
 FT DISULFID 160 177
 FT DISULFID 200 261
 FT CARBOHYD 96 96
 FT CARBOHYD 118 118
 FT CARBOHYD 153 153
 FT CARBOHYD 159 159
 FT CARBOHYD 193 193
 FT CARBOHYD 197 197
 FT CARBOHYD 238 238
 FT CONFLICT 51 51
 FT CONFLICT 121 121

FT CONFLICT 148 148 L -> R (IN REF. 3 AND 4).
 FT CONFLICT 248 248 F -> L (IN REF. 3).
 SQ SEQUENCE 290 AA; 33326 MW; EBD6E156770BD3 CRC64;
 Query Match
 Best Local Similarity 46.2%; Score 37; DB 1; Length 290;
 Matches 9; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
 QY 1 PCVFIR--RVSN 10
 DB 176 PCVFIRKRVIN 187
 Search completed: April 20, 2003, 13:07:36
 Job time: 4.92105 secs

RESULT 2	
Q9LLT2	
ID Q9LLT2	PRELIMINARY; PRT; 367 AA

AC 09LUT2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pollen major allergen 1-2.
 OS Juniperus virginiana (Eastern red cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 NCBI_TaxID=39584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21315424; PubMed=11422137;
 RA Miodoro-Horvut T.M., Goldblum R.M., Brooks E.G.;
 RT "Identification of mutations in the genes for the pollen allergens of
 eastern red cedar (Juniperus virginiana).";
 RL Clin. Exp. Allergy 31:771-778 (2001).
 DR EMBL; AF151427; AAF80166.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF89586 CRC64;

Query Match 76.2%; Score 61; DB 10; Length 367;
 Best Local Similarity 60.0%; Pred. No. 0.006;
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIRKRVSNVILHG 15
 DB 127 PCLFMKRVSHVILHG 141

RESULT 3
 09LUT1 PRELIMINARY; PRT; 367 AA.

ID 09LUT1;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pollen major allergen 1-1.
 OS Juniperus virginiana (Eastern red cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 NCBI_TaxID=39584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21315424; PubMed=11422137;
 RA Miodoro-Horvut T.M., Goldblum R.M., Brooks E.G.;
 RT "Identification of mutations in the genes for the pollen allergens of
 eastern red cedar (Juniperus virginiana).";
 RL Clin. Exp. Allergy 31:771-778 (2001).
 DR EMBL; AF151429; AAF80166.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;

Query Match 76.2%; Score 61; DB 10; Length 367;
 Best Local Similarity 60.0%; Pred. No. 0.006;
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIRKRVSNVILHG 15
 DB 127 PCLFMKRVSHVILHG 141

RESULT 4
 09M4S6 PRELIMINARY; PRT; 367 AA.

ID 09M4S6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.

GN CUP81.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monnaie R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257491; AAF72625.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39835 MW; B4B9C60108C2C5A3 CRC64;

Query Match 76.2%; Score 61; DB 10; Length 367;
 Best Local Similarity 60.0%; Pred. No. 0.006;
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIRKRVSNVILHG 15
 DB 127 PCLFMKRVSHVILHG 141

RESULT 5
 09M4S5 PRELIMINARY; PRT; 367 AA.

ID 09M4S5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUP81.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monnaie R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257492; AAF72626.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0E3263B741 CRC64;

Query Match 76.2%; Score 61; DB 10; Length 367;
 Best Local Similarity 60.0%; Pred. No. 0.006;
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIRKRVSNVILHG 15
 DB 127 PCLFMKRVSHVILHG 141

RESULT 6
 09M4S4 PRELIMINARY; PRT; 367 AA.

ID 09M4S4;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUP81.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 NCBI_TaxID=13469;

[1]
 RN Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RP Monesalve R.I., Villalba M., Rodriguez R.;
 RA "Cloning and expression of Cup s 1, the major allergen of the pollen
 RT of Cupressus sempervirens.";
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257493; AAF72627.1;
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 76.2%; Score 61; DB 10; Length 367;
 Best Local Similarity 60.0%; Pred. No. 0.006;
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIRKRVSNVITNG 15
 DB 127 PCLFMKRVSHVILHG 141

RESULT 7

OQ9M4S3 PRELIMINARY; PRT; 367 AA.
 AC OQ9M4S3.
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monesalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 RL of Cupressus sempervirens.";
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257494; AAF72628.1;
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39832 MW; BSDPBP5A61C07A53 CRC64;

Query Match 76.2%; Score 61; DB 10; Length 367;
 Best Local Similarity 60.0%; Pred. No. 0.006;
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIRKRVSNVITNG 15
 DB 127 PCLFMKRVSHVILHG 141

RESULT 8

OQ9M4S2 PRELIMINARY; PRT; 367 AA.
 AC OQ9M4S2.
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monesalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 RL of Cupressus sempervirens.";

RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257495; AAF72629.1;
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39819 MW; AE7E055A61C07A53 CRC64;

Query Match 76.2%; Score 61; DB 10; Length 367;
 Best Local Similarity 60.0%; Pred. No. 0.006;
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIRKRVSNVITNG 15
 DB 127 PCLFMKRVSHVILHG 141

RESULT 9

OQ93X51 PRELIMINARY; PRT; 367 AA.
 AC OQ93X51.
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Putative allergen jun o 1.
 GN JUN O 1.
 OS Juniperus oxycedrus (Prickly juniper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxID=99008;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=POLLEN;
 RA Iacovacci P., Di Felice G., Pini C.;
 RT "Cloning of Juniperus oxycedrus major allergen.";
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ293767; CAC48400.1;
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase1.
 SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBFD1B9D7 CRC64;

Query Match 76.2%; Score 61; DB 10; Length 367;
 Best Local Similarity 60.0%; Pred. No. 0.006;
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIRKRVSNVITNG 15
 DB 127 PCLFMKRVSHVILHG 141

RESULT 10

O65388 PRELIMINARY; PRT; 390 AA.
 AC O65388.
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE F12F1.22 protein.
 GN F12F1.22.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Vysotskaya V.S., Osborne B.I., Schwartz J.R., Toriumi M., Yu G.,
 RA Kwan A., Oji O., Liu S., Buehler E., Conway A.B., Conway A.R.,
 RA Dewar K., Feng J., Kim C., Kurtz D., Palm C.J., Li Y., Shinn P.,
 RA Sun H., Davis R.W., Becker J.R., Federpfiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F12F1 sequence, complete
 RL Submitted (May-1998) to the EMBL/GenBank/DBJ databases.

RT "The Xanthomonas type III effector protein AvrBs3 modulates plant gene

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; Brassicales; Brassicaceae; Arabidopsis

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; Brassicales; Brassicaceae; Arabidopsis

OK NCB1_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Bevan M., Wedler H., Wambutt R., Bancroft I., Mewes H.W., Mayer K.,

RA Scheller C.; Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Wedler H., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.,

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL022140; CA1811.1; -

DR EMBL; AL161556; CAB79163.1; -

DR InterPro; IPR002022; Amb allergen.

DR InterPro; IPR002048; EF-hand.

DR Pfam; PF00544; pec_lyase; 1.

DR PRINTS; PR00807; AMBALERGEN.

DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.

KW Lyase.

SO SEQUENCE 394 AA; 43476 MW; A48DD586ECF148CA CRC64;

QY 1 PCVFIKRVSNVING 15

Db 150 PCITIQDVNTNVIHG 164

RESULT 15

ID Q43783 PRELIMINARY; PRT; 398 AA.

AC Q43783;

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

PE Pectate lyase (EC 4.2.2.2) (Fragment).

GN PEL.

OS Musa acuminata (Banana).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Zingiberales; Musaceae;

OC Musa.

OX NCB1_TaxID=4641;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=CV. DWARF CAVENDISH; TISSUE=RIPE FRUIT;

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA STRAIN=CV. DWARF CAVENDISH; TISSUE=RIPE FRUIT;

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; X92943; CA63496.1; -

DR HSP; P39116; IEN8.

DR InterPro; IPR002022; Amb allergen.

DR Pfam; PF00544; pec_lyase; 1.

DR PRINTS; PR00807; AMBALERGEN.

KW Lyase.

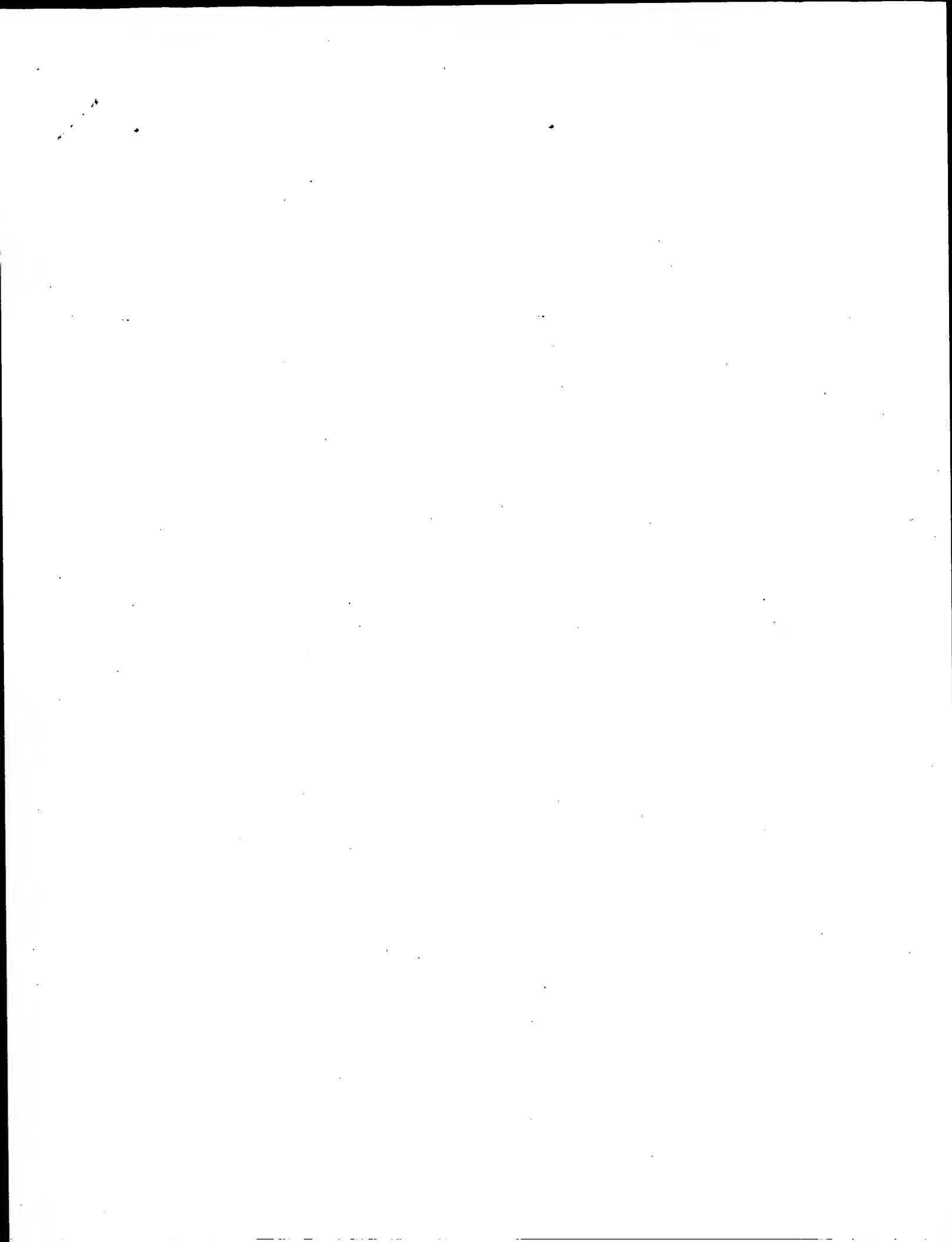
FT NON TER

SO SEQUENCE 398 AA; 43712 MW; A830293E28C6362C CRC64;

QY 1 PCVFIKRVSNVING 15

Db 153 PCITIQDVNTNVIHG 167

Search completed: April 20, 2003, 13:13:06
Job time : 13.6711 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds
(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524D-57
Perfect score: 78
Sequence: 1 KSMKVTVAFNQFGPN 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	78	100.0	374 2 UC2124	major allergen Cry
2	78	100.0	374 2 UC2123	major allergen Cry
3	58	74.4	368 2 G86427	probable pectate 1
4	58	74.4	394 2 T49115	pectate lyase like
5	58	74.4	394 2 T49116	pectate lyase like
6	57	73.1	390 2 H86253	hypothetical prote
7	52	66.7	404 2 S32209	pectate lyase (EC
8	52	66.7	404 2 T05536	pectate lyase (EC
9	51	65.4	455 2 T00856	pectate lyase (EC
10	51	65.4	459 2 G86278	pectate lyase (EC
11	49	62.8	397 2 S26211	hypothetical prote
12	49	62.8	431 2 F86179	pectate lyase (EC
13	49	62.8	450 2 T09524	pectate lyase (EC
14	48	61.5	472 2 T51456	hypothetical prote
15	47	60.3	418 2 T07701	pectate lyase (EC
16	47	60.3	434 2 S29612	pectate lyase (EC
17	47	60.3	438 2 S43335	pectate lyase (EC
18	46	59.0	374 2 T05240	pectate lyase (EC
19	46	59.0	374 2 H85148	pectate lyase (EC
20	46	59.0	398 2 T07058	probable pectate 1
21	46	59.0	463 2 T46165	pectate lyase (EC
22	46	59.0	542 2 T06728	pectate lyase (EC
23	43	55.1	449 2 S27058	pectate lyase (EC
24	42	53.8	397 2 C53240	pectate lyase (EC
25	42	53.8	397 2 C39099	allergen Amb a 1.3
26	42	53.6	1199 2 S77082	pyruvate (Flavodox
27	41	52.6	416 2 S45484	glycerinaldehyde-3-p
28	41	52.6	2833 2 T23064	hypothetical prote
29	41	52.6	2823 2 F87908	protein T22A3.8 [1

30	41	52.6	3102 2	T43291
31	40	51.3	333 2	I84743
32	40	51.3	336 2	I49766
33	40	51.3	363 2	G95237
34	40	51.3	363 2	H98101
35	40	51.3	677 2	B71870
36	39	50.0	210 2	D66531
37	39	50.0	227 2	B90429
38	39	50.0	286 2	T09704
39	39	50.0	348 2	S77104
40	39	50.0	392 2	D53240
41	39	50.0	398 2	B53240
42	39	50.0	398 2	B53240
43	38	48.7	329 2	B39099
44	38	48.7	329 2	G82785
45	38	48.7	364 2	AH1857
			396 2	A39099

ALIGNMENTS

RESULT 1

UC2124 major allergen Cry j I precursor (clone PCCI-15) - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C/Accession: JC2124

R/Source, T: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japan

A/Reference number: JC2123; MUID:94183234; PMID:8135802

A/Accession: JC2124

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-references: GB:026545; NID:9493633; PIDN:BAA05543.1; PID:9493634

A/Experimental source: pollen

A/Note: the authors described carbohydrate binding site for residue 279

C/Superfamily: pectate lyase LAF59

C/Keywords: glycoprotein; pollen

F/1-21/Domain: signal sequence #status predicted <SIG>

F/22-374/Product: major allergen Cry j I (clone PCCI-15) #status predicted <MAT>

F/158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 78; DB 2; Length 374;

Best Local Similarity 100.0%; Pred. No. 8.4e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15

DB 232 KSMKVTVAFNQFGPN 246

RESULT 2

UC2123 major allergen Cry j I precursor (clone PCCI-2-2) - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C/Accession: JC2123; PC2065

R/Source, T: Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japan

A/Reference number: JC2123; MUID:94183234; PMID:8135802

A/Accession: JC2123

A/Molecule type: mRNA

A/Residues: 1-374 <SON>

A/Cross-references: GB:026544; NID:9493631; PIDN:BAA05542.1; PID:9493632

A/Experimental source: pollen

A/Accession: PC2065

A/Molecule type: protein

A/Residues: 22-53;58-81;219-232;236-258;299-307;346-372 <SO>

A/Note: the authors described carbohydrate binding site for residue 279

C/Superfamily: pectate lyase LAF59

C/Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-374/Product: major allergen cys 3 I (clone pCCT-2-2) #status predicted <MAT>
F:158,191,193,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 78; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 8,4e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGR 15
DB 232 KSMKVTVAFNQFGR 246

RESULT 3

G86427 probable pectate lyase [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: G86427

R/Theologian: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: G86427

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-368 <STO>

A/Cross-references: GB:AE005172; NID:g11055821; PIDN:AG28291.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

C/Superfamily: pectate lyase LAT59

Query Match 74.4%; Score 58; DB 2; Length 368;
Best Local Similarity 71.4%; Pred. No. 0.0051;

Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGR 14
DB 229 KSMKVTVAFNQFGR 242

RESULT 4

T49115 pectate lyase like protein - Arabidopsis thaliana

N/Alternate names: protein AT4g22080

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000

C/Accession: T49115

R/Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; M submitted to the Protein Sequence Database, May 2000

A/Reference number: Z25016

A/Accession: T49115

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-394 <BEV>

A/Cross-references: EMBL:AL022140; GSPDB:GN00062; ATSP:AT4g22080

A/Experimental source: cultivar Columbia; BAC clone FIN20

C/Genetics:

A/Map position: 4

A/Intons: 240/3; 275/2; 321/3

C/Superfamily: pectate lyase LAT59

QY 3 MKVTVAFNQFGR 14
DB 256 MKVTVAFNQFGR 267

RESULT 5

T49116 pectate lyase like protein - Arabidopsis thaliana

N/Alternate names: protein AT4g22090

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000

C/Accession: T49116

R/Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; M submitted to the Protein Sequence Database, May 2000

A/Reference number: Z25016

A/Accession: T49116

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-394 <BEV>

A/Cross-references: EMBL:AL022140; GSPDB:GN00062; ATSP:AT4g22090

A/Experimental source: cultivar Columbia; BAC clone FIN20

C/Genetics:

A/Map position: 4

A/Intons: 240/3; 275/2; 321/3

C/Superfamily: pectate lyase LAT59

Query Match 74.4%; Score 58; DB 2; Length 394;
Best Local Similarity 71.7%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 MKVTVAFNQFGR 14
DB 256 MKVTVAFNQFGR 267

RESULT 6

H86253 hypothetical protein [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: H86253

R/Theologian: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: H86253

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-390 <STO>

A/Cross-references: GB:AE005172; NID:g157942; PIDN:AC17625.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

C/Superfamily: pectate lyase LAT59

Query Match 73.1%; Score 57; DB 2; Length 390;
Best Local Similarity 71.4%; Pred. No. 0.0084;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGR 14
DB 244 KSMKVTVAFNQFGR 257

RESULT 7

S12209
pectate lyase (EC 4.2.2.2) - tomato
C/Species: Lycopersicon esculentum (tomato)
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 23-Jul-1999
C/Accession: S12209
R/Budeller, K.A.; Smith, A.G.; Gasser, C.S.
Mol. Gen. Genet. 224, 183-192, 1990
A/Title: Regulation of a stylar transmitting tissue-specific gene in wild-type and trans
A/Reference number: S12209; MUID:91117185; PMID:2277637
A/Accession: S12209
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-404 <BUD>
A/Cross-references: GB:X55193; NID:G19161; PIDN:CAA38979.1; PID:G19162
C/Keywords: carbon-oxygen lyase

Query Match 66.7%; Score 52; DB 2; Length 404;
Best Local Similarity 76.9%; Pred. No. 0.077;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFG 13
|:|:|:|:|:|:|
Db 262 KMQVTVAFNHFG 274

RESULT 8
T05556

pectate lyase (EC 4.2.2.2) P22K18.20 - Arabidopsis thaliana
N/Alternate names: protein P22K18.20
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C/Accession: T05556
R/Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hohnsiegel, J.; Mewes, H.W.; Meyer, K.F.
submitted to the Protein Sequence Database, February 1999
A/Reference number: Z15419
A/Accession: T05556
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-404 <BEV>
A/Cross-references: EMBL:AL035356
A/Experimental source: cultivar Columbia; BAC clone P22K18
C/Genetics:
A/Map position: 4
A/Introns: 37/2; 287/2
A/Note: P22K18.20
C/Superfamily: pectate lyase LAT59
C/Keywords: carbon-oxygen lyase

Query Match 66.7%; Score 52; DB 2; Length 404;
Best Local Similarity 69.2%; Pred. No. 0.077;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFG 13
|:|:|:|:|:|:|
Db 266 KMQVTVAFNHFG 278

RESULT 9
T00856

pectate lyase (EC 4.2.2.2) T20F6.14 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001
C/Accession: T00856; AB4440
R/Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, March 1998
A/Description: Arabidopsis thaliana chromosome II BAC T20F6 genomic sequence.
A/Reference number: Z14206
A/Accession: T00856
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-455 <ROU>
A/Cross-references: EMBL:AC002521; NID:G2947056; PIDN:AAC05350.1; PID:G2947069
A/Experimental source: cultivar Columbia

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, T.; Tallon,
euss, D.; Niekman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: AB4440; MUID:20083487; PMID:10657197
A/Accession: AB4440
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-455 <STO>
A/Cross-references: GB:AE002093; NID:G2947069; PIDN:AAC05350.1; GSPDB:GN00139
C/Genetics:
A/Gene: AC292720
A/Map position: 2
A/Introns: 66/2; 295/3; 376/3
A/Note: T20F6.14
C/Superfamily: pectate lyase LAT59
C/Keywords: carbon-oxygen lyase

Query Match 65.4%; Score 51; DB 2; Length 455;
Best Local Similarity 69.2%; Pred. No. 0.14;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFG 13
|:|:|:|:|:|:|
Db 309 KMQVTVAFNHFG 321

RESULT 10
G86278

hypothetical protein F14L17.19 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 01-Mar-2002
C/Accession: G86278
R/Theologis, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, D.
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Mailli, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: AB6141; MUID:21016719; PMID:11130712
A/Accession: G86278
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-459 <STO>
A/Cross-references: GB:AE005172; NID:G7262684; PIDN:AAF43942.1; GSPDB:GN00141
C/Genetics:
A/Map position: 1
C/Superfamily: pectate lyase LAT59

Query Match 65.4%; Score 51; DB 2; Length 459;
Best Local Similarity 69.2%; Pred. No. 0.14;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFG 13
|:|:|:|:|:|:|
Db 313 KMQVTVAFNHFG 325

RESULT 11
S26211

pectate lyase (EC 4.2.2.2) - common tobacco
C/Species: Nicotiana tabacum (common tobacco)
C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 21-Jul-2000
C/Accession: S26211; S26212; S21933; S22753; S22754
R/Rogers, H.J.; Harvey, A.; Lonsdale, D.M.
Plant Mol. Biol. 20, 493-502, 1992
A/Title: Isolation and characterization of a tobacco gene with homology to pectate lyase
A/Reference number: S26211; MUID:93043039; PMID:1421152

A:Accession: S26211
 A:Molecule type: DNA
 A:Residues: 1-397 <ROG>
 A:Cross-references: EMBL:X67158; NID:g19907; PIND:CAA47630.1; PID:g19908
 A:Accession: S26212
 A:Molecule type: mRNA
 A:Residues: 119-155, 'C', 157-188, 190, 'G', 191-199, 'D', 200, 'R', 203-248, 'N', 250-381 <RO2>
 A:Cross-references: EMBL:X67159; NID:g19909; PIND:CAA47631.1; PID:g3980174
 A:Note: translation of the nucleotide sequence is not complete
 R:Lonsdale, D. M.
 Submitted to the EMBL Data Library, July 1991
 A:Reference number: S21933
 A:Accession: S21933
 A:Molecule type: DNA
 A:Residues: 1-397 <LON>
 A:Cross-references: EMBL:X61102; NID:g19981; PIND:CAA43414.1; PID:g19982
 C:Genetics:
 A:Introns: 193/1; 293/2
 C:Superfamily: pectate lyase LAT59
 C:Keywords: carbon-oxygen lyase

Query Match 62.8%; Score 49; DB 2; Length 397;
 Best Local Similarity 61.5%; Pred. No. 0.28;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFG 13
 ||:|||||
 Db 254 KMQVTVAFNHFG 266

RESULT 12

F86179
 hypothetical protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F86179
 R:Phenology: A.; Becker, J. R.; Palm, C. J.; Federspiel, N. A.; Kaul, S.; White, O.; Alonso, Chin, C. W.; Chung, M. K.; Conn, L.; Conway, A. B.; Conway, A. R.; Creasy, T. H.; Dewar, K.; Jensen, N. F.; Hughes, B.; Hulzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J. L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C. A.; Li, J. H.; Li, Y.; Liu, X.; Liu, S. X.; Liu, Z. A.; Lueros, J. S.; Maill, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzedo, S. L.; Schwartz, J. R.; Shinn, P.; Southwick, A. M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C. M.; Venter, J. C.; Davis, R. W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F86179
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-431 <STO>
 A:Cross-references: GB:AEO05172; NID:g2494113; PIND:AAB80622.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: pectate lyase LAT59

Query Match 62.8%; Score 49; DB 2; Length 431;
 Best Local Similarity 61.5%; Pred. No. 0.31;
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFG 13
 ||:|||||
 Db 289 KMQVTVAFNHFG 301

RESULT 13

T09524
 probable pectate lyase (EC 4.2.2.2) - alfalfa
 C:Species: Medicago sativa (alfalfa)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999
 C:Accession: T09524
 R:Wu, Y.; Qiu, X.; Du, S.; Erickson, L.
 Submitted to the EMBL Data Library, November 1995

A:Reference number: Z16712
 A:Accession: T09524
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-450 <WU>
 A:Cross-references: EMBL:U41472; NID:g1171160; PID:g1171161
 C:Genetics:
 A:Introns: 65/2; 295/3
 C:Function:
 A:Description: catalyzes cleavage of pectate to oligosaccharides
 C:Superfamily: pectate lyase LAT59
 C:Keywords: carbon-oxygen lyase

Query Match 62.8%; Score 49; DB 2; Length 450;
 Best Local Similarity 60.2%; Pred. No. 0.32;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFG 13
 ||:|||||
 Db 309 KMQVTVAFNHFG 321

RESULT 14

T51456
 pectate lyase-like protein - Arabidopsis thaliana
 N:Alternate names: protein F2G14.230
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 17-Nov-2000
 C:Accession: T51456
 R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; N
 Submitted to the Protein Sequence Database, August 2000
 A:Reference number: Z25394
 A:Accession: T51456
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-472 <SAT>
 A:Cross-references: EMBL:AL391146
 A:Experimental source: cultivar Columbia; BAC clone F2G14
 C:Genetics:
 A:Map position: 5
 A:Introns: 65/2; 316/3; 398/3
 A:Note: F2G14.230
 C:Superfamily: pectate lyase LAT59

Query Match 61.5%; Score 48; DB 2; Length 472;
 Best Local Similarity 69.2%; Pred. No. 0.52;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFG 13
 ||:|||||
 Db 330 KMQVTVAFNHFG 342

RESULT 15

T07701
 pectate lyase (EC 4.2.2.2) F17N18.100 - Arabidopsis thaliana
 N:Alternate names: protein F17N18.100
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Jul-1999
 C:Accession: T07701
 R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H. W.; Meyer, K. F. X.; Sch
 Submitted to the Protein Sequence Database, May 1999
 A:Reference number: Z16098
 A:Accession: T07701
 A:Molecule type: DNA
 A:Residues: 1-418 <BEV>
 A:Cross-references: EMBL:AL049751
 A:Experimental source: cultivar Columbia; BAC clone F17N18
 C:Genetics:
 A:Map position: 4
 A:Introns: 26/2; 49/2; 264/3; 299/2; 345/3
 A:Note: F17N18.100
 C:Superfamily: pectate lyase LAT59

Sat May 10 08:44:43 2003

us-09-142-524d-57.rpr

Page 5

C;Keywords: carbon-oxygen lyase

Query Match	60.3%;	Score 47;	DB 2;	Length 418;
Best Local Similarity	61.5%;	Pred NO	0.71;	

Best Local Similarity 61.5%; Pred. No. 0.71;
Matches 8; Conservative 3; Mismatches

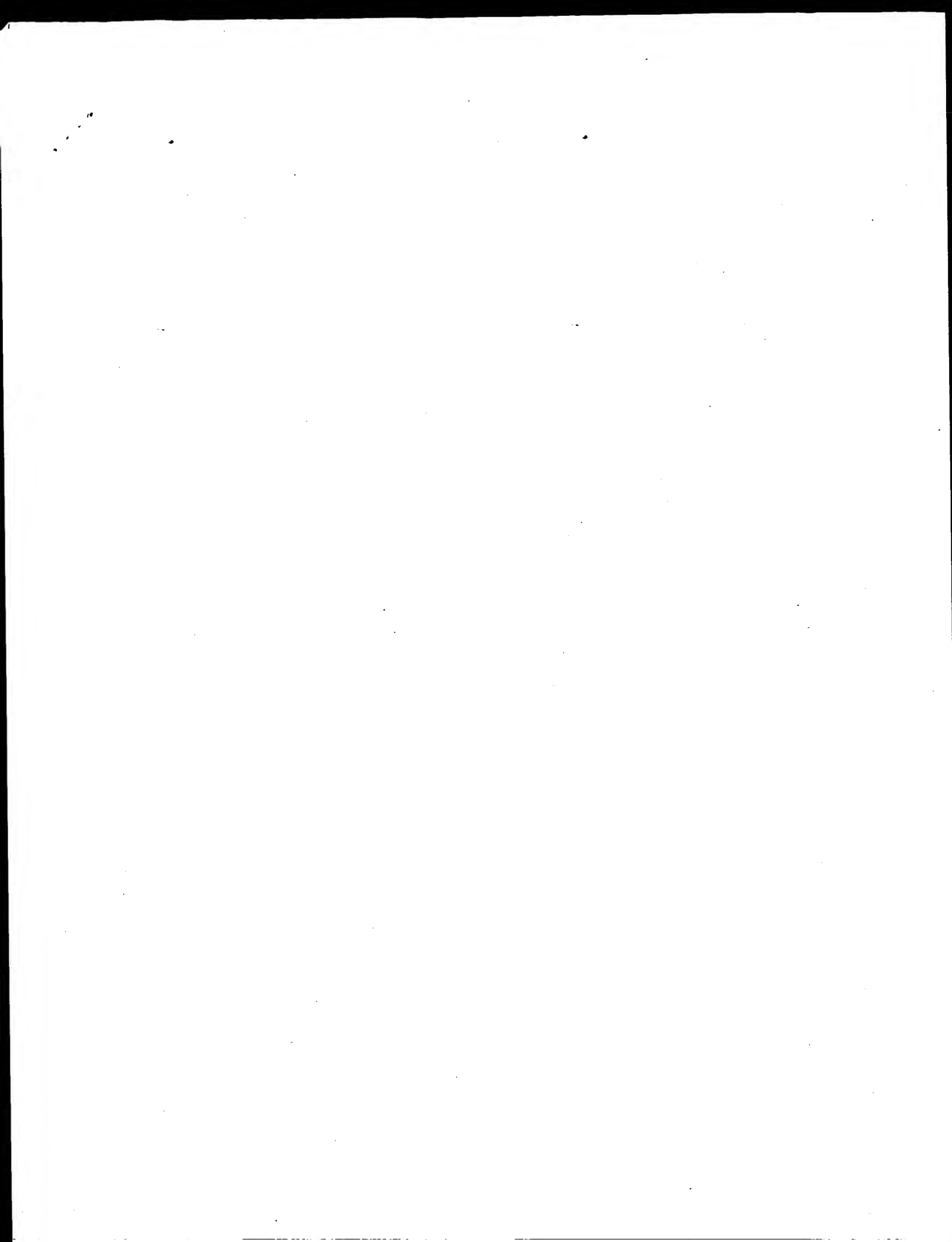
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFG 13

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

Search completed: April 20, 2003, 13:15:46
Job time : 6.07895 secs

Job time : 6.07895 secs



GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524D-57

Perfect score: 78

Sequence: 1 KSMKVTVAENFGPN 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues 112892

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	346	1 MPAL_CUPAR	O96C99 cupressus a
2	78	100.0	367	1 MPAL_JUNAS	P81294 juniperus a
3	78	100.0	374	1 SBR_CRYDA	P18632 cryptomeria
4	78	100.0	375	1 MPAL_CHAOB	O96385 chamaecypar
5	52	66.7	404	1 9612_LYCES	P24396 lycopersico
6	49	62.8	397	1 PEL_TOBAR	P40972 nicotiana t
7	47	60.3	434	1 PEL_LILLO	P40973 lilium long
8	46	59.0	398	1 PES6_LYCES	P15721 lycopersico
9	43	55.1	449	1 PES9_LYCES	P15722 lycopersico
10	42	53.8	397	1 MP13_AMBAR	P27761 ambrosia ar
11	42	53.8	1199	1 NIFU_SYNY3	P52965 synchocyst ar
12	41	52.6	416	1 G3PA_GRAVE	P30724 graciilaria
13	40	51.3	333	1 EFB2_HUMAN	P52799 homo sapien
14	40	51.3	336	1 EFB2_MOUSE	P52799 mus musculu
15	39	50.0	327	1 EFB1_XENIA	O13097 xenopus lae
16	39	50.0	348	1 TRPD_SYNY3	P73617 synchocyst
17	39	50.0	392	1 MP14_AMBAR	P27760 ambrosia ar
18	39	50.0	398	1 MP12_AMBAR	O93932 drosophila ar
19	39	50.0	416	1 O35A_DROME	O73874 brachydanio
20	39	48.7	332	1 EFB2_BRAVE	P27759 ambrosia ar
21	38	48.7	396	1 MP11_AMBAR	P27759 ambrosia ar
22	38	48.7	414	1 G3PA_CHOCR	P30724 graciilaria
23	38	48.7	466	1 YAM8_SCHRO	O10063 schizosacch
24	38	48.7	589	1 TRPG_THEMA	O00708 thermotoga
25	37	47.4	127	1 LYC_COLLI	O02008 drosophila
26	37	47.4	272	1 TRAZ_DROVI	O73612 gallus gall
27	37	47.4	334	1 EFB1_CHICK	O34425 bacillus su
28	37	47.4	340	1 G3P2_BACSU	P52795 mus musculu
29	37	47.4	345	1 EFB1_MOUSE	P52796 ratu mus sapien
30	37	47.4	346	1 EFB1_RAT	P52796 ratu mus sapien
31	37	47.4	346	1 EFB1_HUMAN	O01841 gallus gall
32	37	47.4	367	1 TGM2_CHICK	P47984 felis silve
33	37	47.4	716	1 ZP2_FELCA	

34	36.5	46.8	1126	1 V125_AMVLE	P03589 alfalfa mos
35	36.5	46.8	1248	1 TOFG_SULAC	O08582 sulfolobus
36	36.5	46.8	3075	1 LMAL_HUMAN	P25391 homo sapien
37	36	46.2	258	1 LYCV_BPPI2	P11187 bacterioph
38	36	46.2	258	1 LYCV_BPPI2	P07540 bacterioph
39	36	46.2	331	1 TRPD_BUCSC	O44602 buchnera ap
40	36	46.2	441	1 TRPD_ARATH	O02166 escherichia
41	36	46.2	560	1 YDEN_ECOLI	P77318 escherichia
42	36	46.2	823	1 CNBA_MOUSE	O88502 mus musculu
43	35	44.9	115	1 YE69_HAEIN	P44205 haemophilus
44	35	44.9	173	1 CME_HAEIN	P45036 haemophilus
45	35	44.9	176	1 SLVD_TREPA	O83369 treponema p

ALIGNMENTS

RESULT 1

MPAL_CUPAR STANDARD; PRT; 346 AA.

ID O96C99; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Major pollen allergen Cup a 1.

OS Cupressus arizonica.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferales; Cupressaceae; Cupressus.

NCBI_Taxid=49011;

RP MEDLINE=20571526; PubMed=11122214;

RA Aceituno E., Del Pozo V., Minguez A., Arieta I., Cortegano I.,

RT "Molecular cloning of major allergen from Cupressus arizonica pollen:

RT Cup a 1."

RT Clin. Exp. Allergy 30:1750-1758(2000).

CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.

CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC EMBL: AJ243570; CAB62551.1; -

DR InterPro: IPR002022; Amb allergen.

DR Pfam: PF00544; pec_lyase_1.

DR PRINTS: PR00807; AMBALDERGEN.

KW Allergen; Glycoprotein.

FT CARBOHYD 127 127

FT CARBOHYD 157 157

FT CARBOHYD 272 272

FT CARBOHYD 346 AA; 37589 MW; F1281DCDAIDSDPFD0 CRG64;

SEQUENCE

Query Match. 100.0%; Score 78; DB 1; Length 346;

Best Local Similarity 100.0%; Pred. No. 4.3e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2

MPAL_JUNAS STANDARD; PRT; 367 AA.

ID P81294; 09JUN97;

DT 16-OCT-2001 (Rel. 40, Created)

DE 16-OCT-2001 (Rel. 40, Last sequence update)

FT VARIANT 202 202 S -> T (IN CRY J 1-B).
 FT VARIANT 221 221 L -> S (IN CRY J 1-B).
 FT VARIANT 358 358 Q -> H (IN CRY J 1-B).
 FT VARIANT 361 361 K -> O (IN CRY J 1-B).
 SQ SEQUENCE 374 AA, 40645 MW, 74AB2550248F66F CRC64;

Query Match 100.0%; Score 78; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 4.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTAFAFNQFBN 15
 DB 232 KSMKVTAFAFNQFBN 246

RESULT 4
 MPAL_CHAQB STANDARD; PRT; 375 AA.

AC Q96385; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major pollen allergen Cha o 1 precursor.
 OS Chamaecyparis obtusa (Japanese cypress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Cupressaceae;
 OC Chamaecyparidaceae.
 NCBI_TaxID=13415;
 RN NCBI_TaxID=13415;
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Pollen;
 RX MEDLINE=96265194; PubMed=8676896;
 RA Suzuki M., Komiyama N., Itoh H., Some T., Kuno K., Takagi I.,
 RA Ohta N.;
 RT "Purification, characterization and molecular cloning of Cha o 1, a
 RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen.";
 RL Mol. Immunol. 33:451-460(1996).
 CC -1 SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; D45404; BAA08246.1; -;
 DR InterPro: IPR002022; Amb allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 DR Allergoh; Glycoprotein; Signal.
 FT CHAIN 1 21 MAJOR POLLEN ALLERGEN CHA O 1.
 FT SIGNAL 1 21
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF7066BBF CRC64;

Query Match 100.0%; Score 78; DB 1; Length 375;
 Best Local Similarity 100.0%; Pred. No. 4.7e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTAFAFNQFBN 15
 DB 232 KSMKVTAFAFNQFBN 246

RESULT 5
 9612_LYCSES

ID 9612 LYCSES STANDARD; PRT; 404 AA.
 AC P24395;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Style development-specific protein 9612 precursor.
 GN 9612.

OS Lycopersicon esculentum (tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4081;

RP SEQUENCE FROM N.A.
 RC STRAIN=cv. VF36; TISSUE=Pistil;
 RX MEDLINE=9117185; PubMed=227637;
 RA Budelier K.A., Smith A.G., Gasser C.S.;
 RT "Regulation of a stylar transmitting tissue-specific gene in
 RT wild-type and transgenic tomato and tobacco."
 RL Mol. Gen. Genet. 224:183-192(1990).

CC -1 FUNCTION: MAY HAVE A ROLE IN THE DEVELOPMENT OF THE TRANSMITTING
 CC TISSUE OF THE STYLE AND/OR IN THE EVENTS RELATED TO POLLINATION
 CC SUCH AS SOME ASPECT IN THE FACILITATION OF COMPATIBLE POLLEN TUBE
 CC GROWTH.

CC -1 SUBCELLULAR LOCATION: Secreted.
 CC -1 TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN THE PISTIL WHERE IT IS
 CC FOUND IN THE OUTER FIVE LAYERS OF THE STRANDS OF TRANSMITTING
 CC TISSUE WITHIN THE UPPER TWO-THIRDS OF THE STYLE. FOUND AT MUCH
 CC LOWER LEVELS IN THE ANTHEIRS AND VEGETATIVE ORGANS.
 CC -1 DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND DURING
 CC -1 SIMILARITY: 54% IDENTICAL TO TOMATO PROTEINS P59 (AC P15722)
 CC AND P56 (AC P15721).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X55193; CAA38979.1; -;
 DR PIR; S12209; S12209.
 DR InterPro: IPR002022; Amb allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KW Signal.
 FT CHAIN 1 20 POTENTIAL.
 FT SIGNAL 1 20
 FT CARBOHYD 21 404 STYLE DEVELOPMENT-SPECIFIC PROTEIN 9612.
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 404 AA; 44298 MW; B26ED69B128D8675 CRC64;

Query Match 66.7%; Score 52; DB 1; Length 404;
 Best Local Similarity 76.9%; Pred. No. 0.036;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSMKVTAFAFNQFBN 13
 DB 262 KSMKVTAFAFNQFBN 274

RESULT 6
 PEL_TOBAC STANDARD; PRT; 397 AA.
 AC P40572;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Pectate lyase precursor (EC 4.2.2.2).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=4097;
 RX [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=cv. Samsun; TISSUE=Pollen;
 RX MEDLINE=93043039; PubMed=1421152;
 RA Rogers H.J., Harvey A., Lonsdale D.M.;
 RT "Isolation and characterization of a tobacco gene with homology to
 RT pectate lyase which is specifically expressed during
 RT microsporogenesis";
 RL Plant Mol. Biol. 20:493-502(1992).
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC -1- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED LATE IN POLLEN
 CC DEVELOPMENT.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X67158; CAA47630.1; -;
 DR EMBL; X67159; CAA47631.1; -;
 DR EMBL; X61102; CAA43414.1; -;
 DR PIR; S26211; S26211.
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 DR Lyase; Signal.
 FT CHAIN 1 25
 FT SIGNAL 26 397
 FT ACT_SITE 22 272
 FT CARBOHYD 134 134
 FT CARBOHYD 227 227
 FT CONFLICT 156 156
 FT CONFLICT 189 190
 FT CONFLICT 200 200
 FT CONFLICT 202 202
 FT CONFLICT 249 249
 FT SEQUENCE 397 AA; 44351 MW; EF0A82CE5DA7643F CRC64;
 SQ
 Query Match 62.8%; Score 49; DB 1; Length 397;
 Best Local Similarity 61.5%; Pred. No. 0.13;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 KSMKTVAFNPG 13
 DB 254 KDMKITLVNHPG 266
 ID PEL_LILLO STANDARD; PRT; 434 AA.
 AC P40973;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Pectate lyase precursor (EC 4.2.2.2).
 OS Lilium longiflorum (Trumpet Lily).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 OC Lilium.
 NCBI_TaxID=4690;
 RX [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=cv. Nellie white; TISSUE=Pollen;
 RA Kim S.R., Finkel D.J., An G.;
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z17328; CAA78976.1; -;
 DR EMBL; L18911; AAA33398.1; -;
 DR PIR; S29612; S29612.
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 DR Lyase; Signal.
 FT CHAIN 1 22
 FT SIGNAL 23 434
 FT ACT_SITE 312 312
 FT CARBOHYD 68 68
 FT CARBOHYD 97 97
 FT SEQUENCE 434 AA; 48457 MW; C1F3E30AD2BBD064 CRC64;
 SQ
 Query Match 60.3%; Score 47; DB 1; Length 434;
 Best Local Similarity 81.8%; Pred. No. 0.33;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 3 MKTVAFNPG 13
 DB 296 MQTVAFNHPG 306
 ID PEL_LILLO STANDARD; PRT; 398 AA.
 AC P15721;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Probable pectate lyase P56 precursor (EC 4.2.2.2).
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4081;
 RX [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=cv. VF36; TISSUE=Anther;
 RX MEDLINE=91322485; PubMed=1983191;
 RA Wing R.A., Yamaguchi J., Larebell S.K., Urein V.M., McCormick S.;
 RT "Molecular and genetic characterization of two pollen-expressed genes
 RT that have sequence similarity to pectate lyases of the plant pathogen
 RT Erwinia";
 RL Plant Mol. Biol. 14:117-128(1990).
 CC [2]
 CC REVISIONS.
 CC STRAIN=cv. VF36; TISSUE=Anther;
 CC Wing R.A.;
 CC Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE
 CC GROWTH.
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X15500; CA33524.1; -

DR PIR; S08383; S08383.

DR InterPro; IPR002022; Amb allergen.

DR Pfam; PF00544; pec lyase; 1.

DR PRINTS; PR00807; AMBALLERGEN.

KW Lyase; Multigene family; Signal.

FT SIGNAL 1 27 OR 22 (POTENTIAL).

FT CHAIN 28 398 PROBABLE PEPTATE LYASE P56.

FT ACT SITE 273 273 POTENTIAL.

FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 398 AA; 44563 MW; 8D676250BD8C7C8 CRC64;

Query Match 59.0%; Score 46; DB 1; Length 398;

Best Local Similarity 53.8%; Pred. No. 0.47;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSMKTVAFNORG 13

Db 255 KGMKTLAYNHFG 267

RESULT 9

PE59 LYCES STANDARD; PRT; 449 AA.

AC P15722;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Probable peccate lyase P59 precursor (EC 4.2.2.2).

GN LATS9.

OS Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

OX NCBI_TaxID=4081;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. VF36; TISSUE=Anther;

RA MEDLINE=91322485; PubMed=1983191;

RT Wang R.A., Yamaguchi J., Larabell S.K., Ursin V.M., McCormick S.;

RT "Molecular and genetic characterization of two pollen-expressed genes

RT that have sequence similarity to peccate lyases of the plant pathogen

RT Rwinia.";

RL Plant Mol. Biol. 14:17-28(1990).

CC -1- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE

CC GROWTH.

CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of peccate to give

CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at

CC their non-reducing ends.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.

CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; X15499; CA33523.1; -

DR PIR; S27098; S27098.

DR InterPro; IPR002022; Amb allergen.

DR Pfam; PF00544; pec lyase; 1.

DR PRINTS; PR00807; AMBALLERGEN.

KW Lyase; Multigene family; Signal.

FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 449 PROBABLE PEPTATE LYASE P59.

FT ACT SITE 325 325 POTENTIAL.

FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 449 AA; 50893 MW; 17E3A13F173B03C CRC64;

Query Match 55.1%; Score 43; DB 1; Length 449;

Best Local Similarity 63.6%; Pred. No. 1.9;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 MKVTVAFNORG 13

Db 309 MQITLAFNHFG 319

RESULT 10

MP13 AMBAR STANDARD; PRT; 397 AA.

ID MP13 AMBAR

AC P2761;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Pollen allergen Amb A 1.3 precursor (Antigen E) (Antigen Amb A 1).

OS Ambrosia artemisiifolia (Short ragweed).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;

OC Heliantheae; Ambrosia.

OX NCBI_TaxID=4212;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Pollen;

RA MEDLINE=91093235; PubMed=1702434;

RT Rainier T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,

RT Klapper D.G.;

RT "Cloning of Amb A 1 (antigen E), the major allergen family of short

RT ragweed pollen.";

RL J. Biol. Chem. 266:1229-1236(1991).

RN [2]

RP SEQUENCE FROM N.A., AND VARIANTS.

RC TISSUE=Pollen;

RA MEDLINE=92234570; PubMed=1809687;

RT Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;

RT "Sequence polymorphism of Amb A 1 and Amb A 11, the major allergens

RT in Ambrosia artemisiifolia (short ragweed).";

RT Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).

CC -1- SUBUNIT: MONOMER.

CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.

CC -1- PTM: The N-terminus is blocked.

CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.

CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.

CC AMB A 1/AMB A 11/CRY J 1 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; M62961; AA332668.1; ALT_SEQ.

DR EMBL; M60360; AA332669.1; ALT_SEQ.

DR PIR; C39099; C39099.

DR InterPro; IPR002022; Amb allergen.

DR Pfam; PF00544; pec lyase; 1.

DR PRINTS; PR00807; AMBALLERGEN.

DR Allergen; Allergen; Signal; Multigene family; Polymorphism.

FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 397 POLLEN ALLERGEN AMB A 1.3.

FT VARIANT 48 48 L -> Y.
SQ SEQUENCE 397 AA; 42928 MM; C8DB41257590DD0A CRC64;

Query Match 53.8%; Score 42; DB 1; Length 397;
Best Local Similarity 60.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15
DB 256 KGMATVAFNMFEDN 270

RESULT 11

IFU SYNY3 STANDARD; PRT; 1199 AA.
AC P52965;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-).
GN NIFU OR SL0741.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
CX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N., Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 644 to 924 of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- FUNCTION: OXIDOREDUCTASE REQUIRED FOR THE TRANSFER OF ELECTRONS FROM PYRUVATE TO FLAVODOXIN (POTENTIAL).
CC -1- CATALYTIC ACTIVITY: Pyruvate + CoA + oxidized flavodoxin = acetyl-CoA + CO(2) + reduced flavodoxin.
CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; D64005; BAA10774.1; -
CC InterPro; IPR001450; 4Fe4S_ferredoxin.
CC InterPro; IPR002869; POR_N.
CC InterPro; IPR002880; POR_N.
CC Pfam; PF00037; Fer4; 2.
CC Pfam; PF01558; POR; 1.
CC Pfam; PF01855; POR_N; 1.
CC PROSITE; PS00198; 4FE4S_FERREDOXIN; 2.
CC Oxidoreductase; Nitrogen fixation; Electron transport; Iron-sulfur; 4Fe-4S; Complete proteome.
CC
CC METAL 690
CC METAL 693 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
CC METAL 696 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
CC METAL 700 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
CC METAL 746 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
CC METAL 749 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
CC METAL 752 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
CC METAL 756 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
SQ SEQUENCE 1199 AA; 131457 MM; 12P5C34C3BBD50B6 CRC64;

Query Match 53.8%; Score 42; DB 1; Length 1199;
Best Local Similarity 60.0%; Pred. No. 7.9;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 KSMKVTVAFNQFGPN 15

DB 459 KSGSVTVSHRFGPN 473

RESULT 12
G3PA GRAVE STANDARD; PRT; 416 AA.
ID G3PA GRAVE
AC P30724;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glyceralddehyde 3-phosphate dehydrogenase, chloroplast precursor (EC 1.2.1.13) (NADP-dependent glyceralddehydephosphate dehydrogenase).
GN GAPA.
OS Gracilaria verrucosa.
OC Eukaryota; Rhodophyta; Florideophyceae; Gracilariaceae; Gracilariaceae.
OX NCBI_TaxID=2777;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Norway;
RX MEDLINE=93306755; PubMed=7916671;
RA Zhou Y.H., Ragan M.A.;
RT "CDNA cloning and characterization of the nuclear gene encoding chloroplast glyceralddehyde-3-phosphate dehydrogenase from the marine red alga Gracilaria verrucosa.";
RT Curr. Genet. 23:483-489(1993).
RL [2]
RN [2]
RP REVISIONS.
RX MEDLINE=95042840; PubMed=7954900;
RA Zhou Y.H., Ragan M.A.;
RT "Cloning and characterization of the nuclear gene encoding placid glyceralddehyde-3-phosphate dehydrogenase from the marine red alga Gracilaria verrucosa.";
RT Curr. Genet. 26:79-86(1994).
RL [2]
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NADP(+) = 3-phospho-D-glyceroyl phosphate + NADPH.
CC -1- PATHWAY: Calvin cycle.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; Z15102; CAA78811.1; -
CC EMBL; L22011; AAA33355.1; -
CC PIR; S25596; S25596.
CC PIR; S45484; S45484.
CC HSP; P50362; INMG.
CC InterPro; IPR00173; GAP_dhhydrogenase.
CC InterPro; IPR00044; gpdh; 1.
CC Pfam; PF00280; gpdh_C; 1.
CC PRINTS; PR00076; G3PDHGRNASE.
CC PROSITE; PS00071; GAPDH; 1.
CC Calvin cycle; Oxidoreductase; NADP; Transit peptide; Chloroplast; Multigene family.
CC
CC CHAIN 1 78
CC BINDING 79 416 GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE.
CC ACT_SITE 233 233 GLYCERALDEHYDE 3-PHOSPHATE.
CC ACT_SITE 260 260 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 416 AA; 44337 MM; 33B9437B84230BD8 CRC64;

Query Match 52.6%; Score 41; DB 1; Length 416;
Best Local Similarity 75.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SKMKVAFNPGF 13
DB 78 SKMKVAFNPGF 89

RESULT 13
EPH2_HUMAN STANDARD; PRT; 333 AA.
AC P52759;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ephrin-B2 precursor (EPH-related receptor tyrosine kinase ligand 5)
DE (LEK-5) (HTK ligand) (HTK-L).
GN EPHB2 OR EPLG5 OR LERKS OR HTKL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC- TISSUE=Brain;
RX MEDLINE=96145238; PubMed=8559144;
RA Cerretti D.P., Vanden Bos T., Nelson N., Kozlosky C.J., Reddy P.,
RA Maraskovsky E., Park L.S., Lyman S.D., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Fletcher R.A.,
RT "Isolation of LERK-5: a ligand of the eph-related receptor tyrosine
RT kinases.";
RL Mol. Immunol. 32:1197-1205(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC- TISSUE=Brain;
RX MEDLINE=9519254; PubMed=7534404;
RA Bennett B.D., Zeigler F.C., Gu Q., Fendly B., Goddard A.D.,
RA Gillett N., Matthews W.,
RT "Molecular cloning of a ligand for the EPH-related receptor protein-
RT tyrosine kinase Htk.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:1866-1870(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98192220; PubMed=953549;
RA Vogt T., Stolz W., Welsh J., Jung B., Kerbel R.S., Kobayashi H.,
RA Landthaler M., McClelland M.,
RT "Overexpression of lerk-5/epig5 messenger RNA: a novel marker for
RT increased tumorigenicity and metastatic potential in human malignant
RT melanomas.";
RL Clin. Cancer Res. 4:791-797(1998).
RN [4]
RP FUNCTION: MAY PLAY A ROLE IN CONSTRAINING THE ORIENTATION OF
CC LONGITUDINALLY PROJECTING AXONS (BY SIMILARITY). EPHB4 AND EPHA3.
CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASES EPHB4 AND EPHA3.
CC -1- TISSUE SPECIFICITY: LONG AND KIDNEY.
CC -1- PFM: INDICIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE
CC CYTOPLASMIC DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, U16797; AAA9707.1; -;
DR EMBL, U16797; AAC41752.1; -;
DR EMBL, U81262; AAD03786.1; -;
DR Genew; HGNC:3227; EPHB2.
DR MIM; 600527; -;
DR InterPro; IPRO01799; Ephrin.
DR Pfam; PF00812; Ephrin.1.
DR PRINTS; PR01347; EPHRIN.
DR ProDom; PD002533; Ephrin; 1.

DR PROSITE; PS01299; EPHRIN; 1.
KW Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
FT Signal; Phosphorylation.
FT SIGNAL 1 27
FT CHAIN 28 333
FT DOMAIN 28 229
FT TRANSMEM 230 250
FT DOMAIN 251 333
FT DOMAIN 331 333
FT CARBOHYD 36 36
FT CARBOHYD 139 139
SQ SEQUENCE 333 AA; 36923 MW; 6D9932A632626A9A CRC64;
Query Match 51.3%; Score 40; DB 1; Length 333;
Best local similarity 40.0%; Pred. No. 5.2;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 SKMKVAFNPGF 15
DB 106 QDIKTFKQFSPN 120

RESULT 14
EPH2_MOUSE STANDARD; PRT; 336 AA.
AC P52800;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ephrin-B2 precursor (EPH-related receptor tyrosine kinase ligand 5)
DE (LEK-5) (HTK ligand) (HTK-L) (ELF-2) (ELF-2).
GN EPHB2 OR EPLG5 OR LERKS OR HTKL OR ELF2 OR EPL5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96145238; PubMed=8559144;
RA Cerretti D.P., Vanden Bos T., Nelson N., Kozlosky C.J., Reddy P.,
RA Maraskovsky E., Park L.S., Lyman S.D., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Fletcher R.A.,
RT "Isolation of LERK-5: a ligand of the eph-related receptor tyrosine
RT kinases.";
RL Mol. Immunol. 32:1197-1205(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CB57BL/6J X SJL/J;
RX MEDLINE=9519254; PubMed=7534404;
RA Bennett B.D., Zeigler F.C., Gu Q., Fendly B., Goddard A.D.,
RA Gillett N., Matthews W.,
RT "Molecular cloning of a ligand for the EPH-related receptor protein-
RT tyrosine kinase Htk.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:1866-1870(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=95379837; PubMed=7651410;
RA Bergmann A.D., Cheng H.J., Brambilla R., Klein R., Planagan J.G.,
RT "ELF-2, a new member of the Eph ligand family, is segmentally
RT expressed in mouse embryos in the region of the hindbrain and newly
RT forming somites.";
RL Mol. Cell. Biol. 15:4921-4929(1995).
RN [4]
RP FUNCTION.
RN MEDLINE=20171264; PubMed=10704386;
RA Imondi R., Wideman C., Kaprielian Z.,
RT "Complementary expression of transmembrane ephrins and their receptors
RT in the mouse spinal cord: a possible role in constraining the
RT orientation of longitudinally projecting axons.";
RL Development 127:1397-1410(2000).
CC -1- FUNCTION: MAY PLAY A ROLE IN CONSTRAINING THE ORIENTATION OF
CC LONGITUDINALLY PROJECTING AXONS.

CC -1- SUBUNIT: BINDS TO THE RECEPTOR TYROSINE KINASE EPHB4.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON LATERAL FLOOR PLATE CELLS, SPECIFICALLY ON COMMISSURAL AXON SEGMENTS THAT HAVE PASSED THROUGH THE FLOOR PLATE. EXPRESSED IN CELLS OF THE RETINAL GANGLION CELL LAYER DURING RETINAL AXON GUIDANCE TO THE OPTIC DISC.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE FLOOR PLATE THROUGHOUT THE PERIOD OF COMMISSURAL AXON PATHFINDING.
 CC -1- PFM: INDUCIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE CYTOPLASMIC DOMAIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U16819; AAA9708.1; -;
 DR EMBL: U38847; AAC42052.1; -;
 DR EMBL: U30244; AAA82934.1; -;
 DR MGI: 105097; Efnb2.
 DR InterPro: IPR001799; Ephrin.
 DR Pfam: PF00812; Ephrin_1.
 DR PRINTS: PR01347; EPHRIN.
 DR PRODOM: PD002533; Ephrin; 1.
 DR PROSITE: PS01299; EPHRIN; 1.
 DR Developmental protein; Neurogenesis; Transmembrane; Glycoprotein; Signal; Phosphorylation.
 KM SIGNAL 1 28
 FT CHAIN 29 336
 FT DOMAIN 29 232
 FT TRANSMEM 233 253
 FT DOMAIN 254 336
 FT DOMAIN 334 336
 FT CARBOHYD 39 39
 FT CARBOHYD 142 142
 FT CONFLICT 3 4
 FT CONFLICT 177 177
 FT CONFLICT 177 177
 SQ SEQUENCE 336 AA; 37202 MW; D08894996839554 CRC64;
 Query Match 51.3%; Score 40; DB 1; Length 336;
 Best Local Similarity 40.0%; Pred. No. 5.2;
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 KSMKVTVAFNQPGN 15
 Db 109 QDVKFTIKFOEFSN 123
 RESULT 15
 EPH1_XENLA
 ID EPH1_XENLA STANDARD; PRT; 327 AA.
 AC 013057;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DT Ephrin-B1 precursor (EPH-related receptor tyrosine kinase ligand 2) (IERK-2) (ELK ligand) (ELK-L) (XLERK).
 DE (IERK-2) (ELK ligand) (ELK-L) (XLERK).
 GN EPHB1 OR EPHG2 OR IERK2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97316777; PubMed=9174051;
 RA Jones T.L., Katavanova I., Chong L., Zhou R.P., Daar I.O.;
 RT "Identification of Xlerk, an Eph family ligand regulated during mesoderm induction and neurogenesis in Xenopus laevis.";

BU Oncogene 14:2159-2166 (1997).
 CC -1- FUNCTION: MAY HAVE A ROLE IN THE DEVELOPING MESENCHYMAL AND NERVOUS TISSUE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: IN THE ADULT, EXPRESSED AT LOW LEVELS IN MOST ADULT TISSUES WITH INCREASED LEVELS OBSERVED IN THE KIDNEY, OOCYTES, OVARY AND TESTIS.
 CC -1- PFM: INDUCIBLE PHOSPHORYLATION OF TYROSINE RESIDUES IN THE CYTOPLASMIC DOMAIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EPHRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U31427; AAC35995.1; -;
 DR InterPro: IPR001799; Ephrin.
 DR Pfam: PF00812; Ephrin_1.
 DR PRINTS: PR01347; EPHRIN.
 DR PRODOM: PD002533; Ephrin; 1.
 DR PROSITE: PS01299; EPHRIN; 1.
 DR Developmental protein; Neurogenesis; Transmembrane; Glycoprotein; Signal; Phosphorylation.
 KM SIGNAL 1 20
 FT CHAIN 21 327
 FT DOMAIN 21 225
 FT TRANSMEM 226 246
 FT DOMAIN 247 327
 FT DOMAIN 325 327
 FT CARBOHYD 131 131
 FT CARBOHYD 202 202
 SQ SEQUENCE 327 AA; 36621 MW; 71230CE7F6BE5974 CRC64;
 Query Match 50.0%; Score 39; DB 1; Length 327;
 Best Local Similarity 40.0%; Pred. No. 7.8;
 Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 1 KSMKVTVAFNQPGN 15
 Db 98 KEVRFITKFOEFSN 112
 Search completed: April 20, 2003, 13:07:37
 Job time: 3.92105 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 / Search time 12.6711 Seconds
(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524d-57

Perfect score: 78

Sequence: 1 KSMKVTVAFNQFGPN 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: SPTREMBL 21: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	367	10	Q91LT2
2	78	100.0	367	10	Q91LTI
3	78	100.0	367	10	Q9M4S6
4	78	100.0	367	10	Q9M4S5
5	78	100.0	367	10	Q9M4S4
6	78	100.0	367	10	Q9M4S3
7	78	100.0	367	10	Q9M4S2
8	78	100.0	367	10	Q9M4S1
9	78	100.0	367	10	Q9M4S0
10	78	100.0	367	10	Q9M4S0
11	78	100.0	367	10	Q9M4S0
12	78	100.0	367	10	Q9M4S0
13	78	100.0	367	10	Q9M4S0
14	78	100.0	367	10	Q9M4S0
15	78	100.0	367	10	Q9M4S0
16	78	100.0	367	10	Q9M4S0

17	55	70.5	407	10	Q9SDW4	Q9sdw4 musa acumin
18	52	66.7	226	10	Q23667	Q23667 arabidopsis
19	52	66.7	392	10	Q9FXD8	Q9fxd8 arabidopsis
20	52	66.7	398	10	Q9M505	Q9m505 vitis vinif
21	52	66.7	401	10	Q24554	Q24554 zinnia eleg
22	52	66.7	404	10	Q9S871	Q9s871 arabidopsis
23	52	66.7	408	10	Q9M116	Q9m116 arabidopsis
24	52	66.7	409	10	Q9LT20	Q9lt20 arabidopsis
25	52	66.7	410	10	Q93XJ1	Q93xj1 salix gligi
26	52	66.7	410	10	Q9FMK5	Q9fmk5 arabidopsis
27	52	66.7	432	10	Q93Z25	Q93z25 arabidopsis
28	51	65.4	324	10	Q8S345	Q8s345 capsicum an
29	51	65.4	349	10	Q93Z77	Q93z77 arabidopsis
30	51	65.4	455	10	Q64510	Q64510 arabidopsis
31	51	65.4	459	10	Q23665	Q23665 arabidopsis
32	51	65.4	459	10	Q9M9S2	Q9m9s2 arabidopsis
33	50	64.1	408	10	Q9CSM8	Q9csm8 arabidopsis
34	49	62.8	263	10	Q04084	Q04084 arabidopsis
35	49	62.8	354	10	Q94001	Q94001 arabidopsis
36	49	62.8	431	10	Q23017	Q23017 arabidopsis
37	49	62.8	450	10	Q40319	Q40319 medicago sa
38	49	62.8	475	10	Q9S8R4	Q9s8r4 arabidopsis
39	48	61.5	143	10	Q94B14	Q94b14 vitis vinif
40	48	61.5	368	10	Q94FT5	Q94ft5 fragaria an
41	48	61.5	392	10	Q9FM66	Q9fm66 arabidopsis
42	48	61.5	405	10	Q24416	Q24416 fragaria an
43	48	61.5	405	10	Q94FT6	Q94ft6 fragaria an
44	48	61.5	452	10	Q9LRM5	Q9lrm5 arabidopsis
45	48	61.5	472	10	Q9LFP5	Q9lfp5 arabidopsis

ALIGNMENTS

RESULT 1

ID Q91LT2 PRELIMINARY; PRT; 367 AA.
AC Q91LT2;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Pollen major allergen 1-2.
OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX NCBI_TaxID=39584;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2135424; PubMed=11422137;
RA Midoro-Horiuchi T.M., Goldblum R.M., Brooke E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778(2001).
DR EMBL, AF15427; AAF80164.1; -
DR InterPro: IPR02022; Amb_allergen.
DR Pfam: PF00544; pec lyase; 1.
DR PRINTS: PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF89586 CRC64;

Query Match 100.0%; Score 78; DB 10; Length 367;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSMKVTVAFNQFGPN 15
Db 232 KSMKVTVAFNQFGPN 246

RESULT 2
Q91LTI PRELIMINARY; PRT; 367 AA.
ID Q91LTI;
AC Q91LTI;
DT 01-OCT-2000 (TREMREL. 15, Created)

01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Pollen major allergen 1.1.
 OS Juniperus virginiana (Eastern red cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxID=39584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21315424; PubMed=11422137;
 RA Miodoro-Horvut T.M., Goldblum R.M., Brooks E.G.;
 RT "Identification of mutations in the genes for the pollen allergens of
 eastern red cedar (Juniperus virginiana).";
 RL Clin. Exp. Allergy 31:771-778 (2001).
 DR EMBL: AF151429; AAF80166.1; -;
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR PRINTS: PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74B4711 CRC64;
 Query Match 100.0%; Score 78; DB 10; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15
 DB 232 KSMKVTVAFNQFGPN 246

RESULT 3
 QM4S6 PRELIMINARY; PRT; 367 AA.

ID QM4S6
 AC QM4S6;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF257491; AAF72625.1; -;
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR PRINTS: PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39835 MW; BAE9C60108C2C5A3 CRC64;
 Query Match 100.0%; Score 78; DB 10; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15
 DB 232 KSMKVTVAFNQFGPN 246

RESULT 4
 QM4S5 PRELIMINARY; PRT; 367 AA.

ID QM4S5
 AC QM4S5;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF257492; AAF72626.1; -;
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR PRINTS: PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0E3263B741 CRC64;
 Query Match 100.0%; Score 78; DB 10; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15
 DB 232 KSMKVTVAFNQFGPN 246

RESULT 5
 QM4S4 PRELIMINARY; PRT; 367 AA.

ID QM4S4
 AC QM4S4;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF257493; AAF72627.1; -;
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR PRINTS: PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74B3D6E CRC64;
 Query Match 100.0%; Score 78; DB 10; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15
 DB 232 KSMKVTVAFNQFGPN 246

RESULT 6
 QM4S3 PRELIMINARY; PRT; 367 AA.

ID QM4S3
 AC QM4S3;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257494; AAF72628.1; -
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; Pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39832 MW; B5DF5A61C07A53 CRC64;

Query Match 100.0%; Score 78; DB 10; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15
 DB 232 KSMKVTVAFNQFGPN 246

RESULT 7

Q9M4S2 PRELIMINARY; PRT; 367 AA.
 ID Q9M4S2
 AC Q9M4S2;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUP81.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;

RA "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257495; AAF72629.1; -
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam; PF00544; Pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39819 MW; AE7E055A61C07A53 CRC64;

Query Match 100.0%; Score 78; DB 10; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15
 DB 232 KSMKVTVAFNQFGPN 246

RESULT 8

Q93X51 PRELIMINARY; PRT; 367 AA.
 ID Q93X51
 AC Q93X51;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Putative allergen jun o 1.
 GN JUN O 1.
 OS Juniperus oxycedrus (Prickly juniper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxID=99008;

RA "Cloning of Juniperus oxycedrus major allergen.";
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ293767; CAC48400.1; -

DR InterPro: IPR002022; Amb_allergen.
 DR Pfam; PF00544; Pec_lyase; 1.
 SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBFD1B9D7 CRC64;

Query Match 100.0%; Score 78; DB 10; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15
 DB 232 KSMKVTVAFNQFGPN 246

RESULT 9

Q8RUR1 PRELIMINARY; PRT; 374 AA.
 ID Q8RUR1
 AC Q8RUR1;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Cry j 1 precursor.
 GN Cry j 1.1 OR Cry j 1.2.
 OS Cryptomeria japonica (Japanese cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
 OX NCBI_TaxID=3369;

RA "Isolation and characterization of cDNAs encoding major allergen Cry j
 1 from Cryptomeria japonica pollen.";
 RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB081309; BAB86286.1; -
 DR EMBL; AB081310; BAB86287.1; -
 FT SIGNAL.
 FT CHAIN 1 21 POTENTIAL.
 SQ SEQUENCE 374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;

Query Match 100.0%; Score 78; DB 10; Length 374;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15
 DB 232 KSMKVTVAFNQFGPN 246

RESULT 10

Q93XL6 PRELIMINARY; PRT; 367 AA.
 ID Q93XL6
 AC Q93XL6;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative allergen Cup a 1 precursor.
 GN CUP A 1.
 OS Cupressus arizonica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=49011;

RA "Cloning of Cupressus arizonica major allergen.";
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ278498; CAC37790.2; -
 KW SIGNAL.
 FT CHAIN 1 21 POTENTIAL.
 SQ SEQUENCE 367 AA; 39809 MW; AAF97260423A9F28 CRC64;

Query Match 93.6%; Score 73; DB 10; Length 367;
 Best Local Similarity 100.0%; Pred. No. 3.3e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SMKVTVAFNQFGPN 15
 DB 233 SMKVTVAFNQFGPN 246

RESULT 11
 Q9FY87 PRELIMINARY; PRT; 297 AA.
 AC Q9FY87
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Major pollen allergen-like protein.
 GN T5E8 80.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL391712; CAC05454.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 SQ SEQUENCE 297 AA; 32971 MW; A75287A90A96D240 CRC64;

Query Match 79.5%; Score 62; DB 10; Length 297;
 Best Local Similarity 73.3%; Pred. No. 0.0029;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15
 DB 181 KSMKVTVAFNQFGPN 195

RESULT 12
 Q9C8G4 PRELIMINARY; PRT; 368 AA.
 AC Q9C8G4
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Pectate lyase, putative.
 GN T4K22.5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CV, COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buglner E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gall J.E., Goldenfild A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
 RA Hunter T.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,

RA Kim C.U., Koo H.L., Kremenskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Malt R., Marzalli A.,
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizko M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.O., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utecherback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:816-820(2000).
 DR EMBL; AC025295; AAG51103.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KW Lyase.
 SQ SEQUENCE 368 AA; 40576 MW; F35EDED82CED34BC CRC64;

Query Match 74.4%; Score 58; DB 10; Length 368;
 Best Local Similarity 71.4%; Pred. No. 0.02;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 14
 DB 229 KSMKVTVAFNQFGPN 242

RESULT 13
 ID 065457 PRELIMINARY; PRT; 394 AA.
 AC 065457
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Pectate lyase like protein.
 GN F1R20.190 OR AT4G22090.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Wedler H., Wambutt R., Bancroft I., Mewes H.W., Mayer K.,
 RA Schueller C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wedler H., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL022140; CAA18112.1; -
 DR EMBL; AL161556; CAB79164.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KW Lyase.
 SQ SEQUENCE 394 AA; 43299 MW; 3912C1F6A8C6E3F5 CRC64;

Query Match 74.4%; Score 58; DB 10; Length 394;
 Best Local Similarity 91.7%; Pred. No. 0.021;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 MKVTVAFNQFGPN 14
 DB 256 MKVTVAFNQFGPN 267

RESULT 14

```

065456 ID 065456 PRELIMINARY; PRT; 394 AA.
AC 065456.
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Pectate lyase like protein.
GN F12P1.280 OR AT4G22080.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Wedler H., Mambutt R., Bancroft I., Mewes H.W., Mayer K.,
RA Schueller C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Wedler H., Mambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022140; CA1811.1; -.
DR EMBL; AL161556; CAB79163.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
KW lyase.
SQ SEQUENCE 394 AA; 43476 MW; A48DD586ECF148CA CRC64;

Query Match
Best Local Similarity 74.4%; Score 58; DB 10; Length 394;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 MKVTVAFNQFGP 14
Db 256 MKVTVAFNHFGP 267

RESULT 15
O65388 ID 065388 PRELIMINARY; PRT; 390 AA.
AC 065388.
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE F12P1.22 protein.
GN F12P1.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. COLUMBIA;
RA Vysotskaya V.S., Osborne B.I., Schwartz J.R., Toriumi M., Yu G.,
RA Kwan A., Oji O., Liu S., Buehler E., Conway A.B., Conway A.R.,
RA Dewar K., Feng J., Kim C., Kurtz D., Palm C.J., Li X., Shinn P.,
RA Sun H., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F12P1 sequence, complete
sequence."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002131; AAC17625.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.

```

```

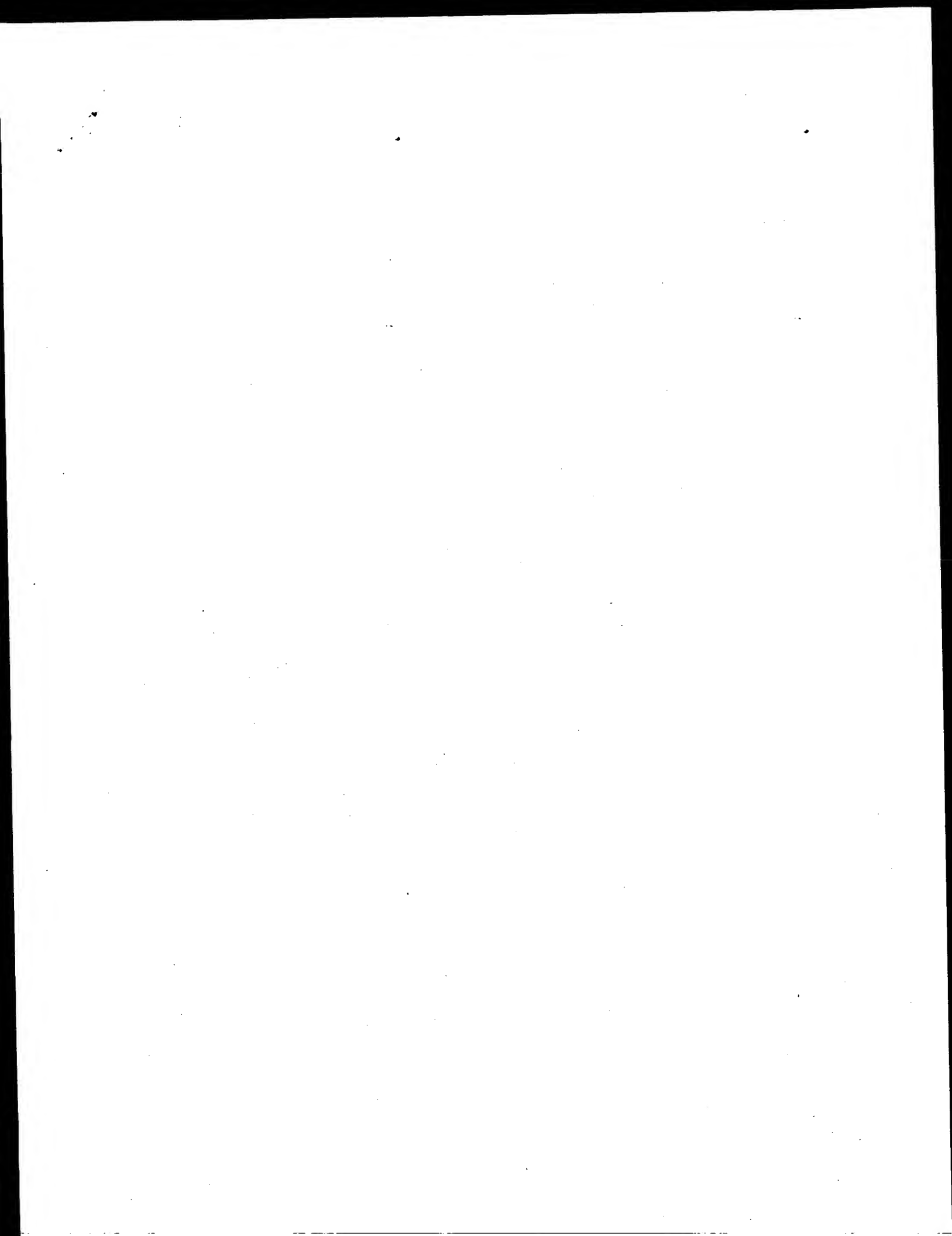
SQ SEQUENCE 390 AA; 43354 MW; 7E4D3CA9BB398CB7 CRC64;

Query Match
Best Local Similarity 73.1%; Score 57; DB 10; Length 390;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGP 14
Db 244 KMKVTVAFNHFGP 257

Search completed: April 20, 2003, 13:13:07
Job time : 13.6711 secs

```



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds
(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524d-97

Perfect score: 81

Sequence: 1 VDGIIAAYQNPASMK 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	514	2	S48730
2	81	100.0	514	2	JC2498
3	61	75.3	507	2	JC7366
4	48	59.3	514	2	JC7100
5	45	55.6	374	2	T02021
6	45	55.6	1113	2	T47381
7	45	55.6	1175	2	D85089
8	45	55.6	1198	2	B86402
9	45	55.6	1201	2	F86386
10	45	55.6	1285	2	H85041
11	45	55.6	1305	2	H86559
12	45	55.6	1314	2	T47331
13	45	55.6	1756	2	T02599
14	44	54.3	521	2	T37252
15	43	53.1	646	2	B96971
16	43	53.1	1036	2	AP2338
17	43	53.1	1102	2	AB1940
18	43	53.1	1102	2	AB1940
19	43	53.1	1645	2	AG1897
20	42	51.9	1207	1	BWASBR
21	42	51.9	1108	2	T33605
22	42	51.9	1140	2	AD3032
23	42	51.9	1140	2	H98253
24	41	50.6	327	2	AE2173
25	41	50.6	327	2	A70435
26	41	50.6	465	2	AC0659
27	40.5	50.0	544	2	PMRT
28	40	49.4	94	2	T24633
29	40	49.4	94	2	E91292
					H86133

Cry j II protein - second major aller Jun a 2 protein - polygalacturonase hypothetical prote hypothetical prote hypothetical prote hypothetical prote hypothetical prote hypothetical prote probable matrix me setine protein kin two-component sens methyl-accepting c two-component hybr bling protein - Eme hypothetical prote conserved hypothet hypothetical prote cysteine synthase hypothetical prote probable ATP-bind phosphoglucosylase hypothetical prote hypothetical prote

30	40	49.4	200	2	D72389
31	40	49.4	212	2	J01378
32	40	49.4	230	2	E70927
33	40	49.4	249	2	S75749
34	40	49.4	302	2	T40226
35	40	49.4	343	2	E84129
36	40	49.4	351	1	SAVZAV
37	40	49.4	351	1	SAVZAV
38	40	49.4	353	1	SAVZVC
39	40	49.4	422	2	B75501
40	40	49.4	644	2	F82145
41	40	49.4	2348	2	AD1841
42	39.5	48.8	573	2	T23589
43	39	48.1	126	1	FAD01
44	39	48.1	127	2	B81799
45	39	48.1	127	2	C81069

ALIGNMENTS

RESULT 1
S48730
Cry j II protein - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C:Accession: S48730
R:Namba, M.; Kurose, M.; Toriige, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; FERS Lett. 353, 124-128, 1994
A>Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar
A:Reference number: S48730; MUID:95010777; PMID:7926035
A:Accession: S48730
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-514 <NM>
A:Cross-References: GB:D37765; NID:9577695; PIDN:BAA07021.1; PID:dl007598; PID:9577696

Query Match 100.0%; Score 81; DB 2; Length 514;
Best local similarity 100.0%; Pred. No. 3.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGIIAAYQNPASMK 15
DB 120 VDGIIAAYQNPASMK 134

RESULT 2
JC2498
second major allergen Cry j II precursor - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C:Accession: JC2498; PC2346; A60147
R:Komiyama, N.; Some, T.; Shimizu, K.; Morikubo, K.; Kino, K. Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
A>Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese cedar
A:Reference number: JC2498; MUID:94271186; PMID:8002972
A:Accession: JC2498
A:Molecule type: mRNA
A:Residues: 1-514 <NM>
A:Cross-References: DDBJ:D29772; NID:9506857; PIDN:BAA06172.1; PID:9506858
A:Accession: PC2346
A:Molecule type: protein
A:Residues: 52-61 <KO2>
R:Sakaguchi, M.; Inoue, S.; Tanai, M.; Ando, S.; Ueki, M.; Matuhasi, T. Allergy 45, 309-312, 1990
A>Title: Identification of the second major allergen of Japanese cedar pollen.
A:Reference number: A60147; MUID:90342988; PMID:2382797
A:Accession: A60147
A:Molecule type: protein
A:Residues: 55-64 <SAK>
C:Keywords: glycoprotein, pollen
F1-54/Domain: signal sequence #status predicted <SIG>
F155-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 81; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGIIAAYONPASWK 15
DB 120 VDGIIAAYONPASWK 134

RESULT 3

UC7366

Jun a 2 proteain - mountain cedar

C:Species: Juniperus ashei (mountain cedar)

C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000

C:Accession: UC7366; PC7093

R:Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.

Biochem. Biophys. Res. Commun. 275, 195-202, 2000

A:Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all

A:Reference number: UC7366

A:Accession: UC7366

A:Molecule type: mRNA

A:Residues: 1-507 <YOK>

A:Cross-references: GB:AJ404653

A:Accession: PC7093

A:Molecule type: protein

A:Residues: 55-63 <YOK>

C:Comment: This protein, a second major allergen of mountain cedar pollen, which is invo

o the polygalacturonase family.

C:Keywords: glycoprotein; pollen

Query Match 75.3%; Score 61; DB 2; Length 507;
Best Local Similarity 73.3%; Pred. No. 0.012;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VDGIIAAYONPASWK 15
DB 121 VDGIIAAYONPASWK 135

RESULT 4

UC7100

polygalacturonase Cha o 2 - Japanese cypress

C:Species: Chamaecyparis obtusa (Japanese cypress)

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000

C:Accession: UC7100; PC7026

R:Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.

Biochem. Biophys. Res. Commun. 263, 166-171, 1999

A:Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all

A:Reference number: UC7100; MUID:99417540; PMID:10486272

A:Accession: UC7100

A:Molecule type: mRNA

A:Residues: 1-514 <MOR>

A:Accession: PC7026

A:Molecule type: protein

A:Residues: 51-62 <MOR>

Query Match 59.3%; Score 48; DB 2; Length 514;
Best Local Similarity 60.0%; Pred. No. 2.3;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VDGIIAAYONPASWK 15
DB 120 VDGIIAAYONPASWK 134

RESULT 5

T02021

hypothetical protein T9E19.2 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999

C:Accession: T02021

R:Stromatt, C.; Johnson, D.; Le, T.
submitted to the EMBL Data Library, November 1998

A:Description: The sequence of A. thaliana T9E19.

A:Reference number: Z14496

A:Accession: T02021

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-374 <STR>

A:Cross-references: EMBL:AF104920; NID:g3859610; PID:g3859611

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Introns: 279/3

A>Note: T9E19.2

Query Match 55.6%; Score 45; DB 2; Length 374;
Best Local Similarity 63.6%; Pred. No. 5.6;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAAYONP 11
DB 154 VDGIIAAYONP 164

RESULT 6

T47381

hypothetical protein T5C2.90 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47381

R:Obermaier, B.; Ottenwajlder, B.; Duchemin, D.; Zeitler, K.; Mewes, R.W.; Lemcke, K.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24463

A:Accession: T47381

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1113 <OBE>

A:Cross-references: EMBL:AL138664

A:Experimental source: cultivar Columbia, BAC clone T5C2

C:Genetics:

A:Map position: 3

A:Introns: 42/1; 351/3; 418/1; 617/3; 659/3; 690/3; 732/3; 767/3; 801/3; 816/3; 845/3

A>Note: T5C2.90

Query Match 55.6%; Score 45; DB 2; Length 1113;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAAYONP 11
DB 226 VDGIIAAYONP 236

RESULT 7

D85089

hypothetical protein AT4g0880 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

C:Accession: D85089

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp

ing Island, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: D85089

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1175 <STO>

A:Cross-references: GB:NC_001268; NID:g7267530; PIDN:CAB78012.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g08880

A:Map position: 4

Query Match 55.6%; Score 45; DB 2; Length 1175;

Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAAYONP 11
||||:|
Db 90 VDGVAHAHNP 100

RESULT 8

hypothetical protein F28U5.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: E86402
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86402
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1198 <STO>
A:Cross-references: GB:AE005172; NID:G1098930; PIDN:AMG26070.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAAYONP 11
||||:|
Db 226 VDGVAHAHNP 236

RESULT 9

hypothetical protein F14G11.4 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: F86386
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86386
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1201 <STO>
A:Cross-references: GB:AE005172; NID:G11560181; PIDN:AMG38123.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAAYONP 11
||||:|

Db 226 VDGVAHAHNP 236

RESULT 10

hypothetical protein AT4G03300 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: H85041
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory
Nature 402, 769-777, 1999
A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85041
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1385 <STO>
A:Cross-references: GB:NC_001268; NID:G7270200; PIDN:CA877815.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4G03300
A:Map position: 4

Query Match
Best Local Similarity 63.6%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAAYONP 11
||||:|
Db 378 VDGVAHAHNP 388

RESULT 11

hypothetical protein F5F19.8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96559
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H96559
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1305 <STO>
A:Cross-references: GB:AE005173; NID:G4220449; PIDN:ADD12676.1; GSPDB:GN00141
C:Genetics:
A:Gene: F5F19.8
A:Map position: 1

Query Match
Best Local Similarity 63.6%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAAYONP 11
||||:|
Db 226 VDGVAHAHNP 236

RESULT 12

hypothetical protein F7P3.20 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47331
R:Vitale, D.; Lignori, R.; Flores, M.; Argirou, A.; De Simone, V.; Mewes, H.W.; Rudd,

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24461
A:Accession: T47331
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1314 <VIT>
A:Cross-references: EMBL:AL138663
A:Experimental source: Cultivar Columbia; BAC clone F7P3
C:Genetics:
A:Map position: 3
A:Intons: 42/1; 351/3; 416/1; 615/3; 657/3; 712/3; 754/3; 789/3; 819/3; 839/3; 873/3; 9
A:Note: F7P3.20

Query Match 55.6%; Score 45; DB 2; Length 1314;
Best Local Similarity 63.6%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAAYONP 11
||:|:|:|
Db 226 VDGIVAAHNSNP 236

RESULT 13

T02599
hypothetical protein At2g14770 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F26C24.9
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02599; C84521
R:Rounisley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.
A:Reference number: Z14660
A:Accession: T02599
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1756 <ROU>
A:Cross-references: EMBL:AC004705; NID:G3252804; PID:G3252818
A:Experimental source: Cultivar Columbia
R:Lin, X.; Kaul, S.; Rounisley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.
euse, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: C84521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1756 <STO>
A:Cross-references: GB:AE002093; NID:G3252818; PID:AA024188.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g14770; F26C24.9
A:Map position: 2
A:Intons: 42/1; 351/3; 418/1; 617/3; 659/3; 690/3; 725/1; 758/3; 778/3; 812/3; 902/2; 9

Query Match 55.6%; Score 45; DB 2; Length 1756;
Best Local Similarity 63.6%; Pred. No. 28;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAAYONP 11
||:|:|:|
Db 226 VDGIVAAHNSNP 236

RESULT 14

T37252
probable matrix metalloproteinase (EC 3.4.24.-) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T37252
R:Wada, K.; Sato, H.; Kinoh, H.; Kajita, M.; Yamamoto, H.; Seiki, M.
Gene 211, 57-62, 1998
A:Title: Cloning of three Caenorhabditis elegans genes potentially encoding novel matrix

A:Reference number: Z21645; MUID:98241501; PMID:9573338
A:Accession: T37252
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-521 <MAD>
A:Cross-references: EMBL:AB007817; NID:G3152405; PIDN:BA28353.1; PID:G3152406
A:Experimental source: strain N2
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprote
C:Keywords: hydrolase; metalloproteinase

Query Match 54.3%; Score 44; DB 2; Length 521;
Best Local Similarity 40.0%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 VDGIIAAYONPASPW 15
||:|:|:|
Db 229 IDSVMPAYDTPRCWK 243

RESULT 15

B96971
serine protein kinase (prkA protein), p-loop containing [imported] - Clostridium acet
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: B96971
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Le
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B96971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-646 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78557.1; PID:G15023447; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0579
A:Superfamily: Bacillus protein kinase A

Query Match 53.1%; Score 43; DB 2; Length 646;
Best Local Similarity 53.8%; Pred. No. 22;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 DGIILAYONPASW 14
||:|:|:|
Db 301 DGIILAHNSNEAW 313

Search completed: April 20, 2003, 13:15:49
Job time : 9.07895 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 seconds
(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524D-97

Sequence: 1 VDGIIAAYGNPASWK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	514	1	MPA2_CRYUA
2	45	55.6	373	1	ACOC_ALCEU
3	43	53.1	2073	1	BIME_EMENT
4	41	50.6	327	1	CYSM_AQUAE
5	41	50.6	561	1	PGMU_RAT
6	40.5	50.0	544	1	YRT3_CABEL
7	40	49.4	230	1	TRMD_MYCTU
8	40	49.4	302	1	RE14_SCHPO
9	40	49.4	351	1	VB19_VACCD
10	40	49.4	351	1	VB19_VACCV
11	40	49.4	353	1	VB19_VACCC
12	40	49.4	444	1	GLGA_DEIRA
13	39	48.1	126	1	PRO1_DICDI
14	39	48.1	299	1	CARB_STRTH
15	39	48.1	356	1	GPDA_TREDA
16	39	48.1	561	1	PGMU_HUMAN
17	39	48.1	561	1	PGMU_MOUSE
18	39	48.1	561	1	PGMU_RABIT
19	39	48.1	3461	1	RELN_MOUSE
20	38.5	47.5	200	1	CYDX_RAT
21	38.5	47.5	327	1	YJBN_HAELN
22	38	46.9	186	1	YER7_HAELN
23	38	46.9	227	1	GYRA_SPTIT
24	38	46.9	306	1	RAM2_CANML
25	38	46.9	440	1	SECY_CORGL
26	38	46.9	476	1	SECY_ARATH
27	38	46.9	566	1	MEHM_ECOLI
28	38	46.9	775	1	MGDI_MOUSE
29	38	46.9	775	1	MGDI_RAT
30	38	46.9	827	1	RED1_YEAST
31	38	46.9	1163	1	ITAX_HUMAN
32	38	46.9	1582	1	TU30_RALUSO
33	38	46.9	3462	1	RELN_RAT

34	37	45.7	209	1	YE28_AQUAE	067421 aquifex aeo
35	37	45.7	351	1	VB19_VARV	P37395 variola vir
36	37	45.7	362	1	CNTR_CHICK	P51641 gallus gall
37	37	45.7	437	1	VIPS_MOUSE	P41588 mus musculu
38	37	45.7	593	1	US22_HUMVA	P09722 human cytom
39	37	45.7	605	1	PHAC_METEX	P52070 methylolact
40	37	45.7	644	1	YEAG_ECOLI	P77391 escherichia
41	37	45.7	778	1	MGDI_HUMAN	O955V3 homo sapien
42	37	45.7	821	1	GYRA_BACSU	P05653 bacillus su
43	37	45.7	835	1	BGAL_LYCES	P48980 lycopersico
44	37	45.7	1045	1	PRTS_SERMA	P09489 serratia ma
45	36.5	45.1	500	1	CPDG_CAVVO	O64403 cavia porce

ALIGNMENTS

RESULT 1	ID	MPA2_CRYUA	STANDARD	PRT	514 AA.
AC	P43212				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)				
DE	(Major pollen allergen Cry j 2) (Cry j II).				
OS	Cryptomeria japonica (Japanese cedar).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.				
OX	NCBI_Taxid=3369;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE=Pollen;				
RA	MEDLINE=95010777; Pubmed=7926035;				
RA	Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,				
RT	Usui M., Kurimoto M.;				
RT	"Molecular cloning of the second major allergen, Cry j II, from				
RL	Japanese cedar pollen.";				
RL	FEBS Lett. 353:124-128(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Pollen;				
RA	MEDLINE=94271186; Pubmed=8002972;				
RA	Komiyama N., Sone T., Shimizu K., Morikubo K., Kano K.;				
RT	"cDNA cloning and expression of Cry j II the second major allergen of				
RL	Japanese cedar pollen.";				
RL	Biochem. Biophys. Res. Commun. 201:1021-1028(1994).				
RN	[3]				
RP	SEQUENCE OF 55-64.				
RA	MEDLINE=90342988; Pubmed=2382797;				
RA	Sakaguchi M., Inouye S., Taniat M., Ando S., Usui M., Matuhasi T.;				
RT	"Identification of the second major allergen of Japanese cedar				
RL	pollen.";				
RL	Allergy 45:309-312(1990).				
CC	-1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-				
CC	galactosiduronic linkages in pectate and other galacturonans.				
CC	-1- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).				
CC	-1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES				
CC	(POLYGALACTURONASES).				
CC	-----				
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; D37765; BAA07021.1; -				
DR	EMBL; D29772; BAA06172.1; -				
DR	HSSP; P26509; 1BHE.				
DR	InterPro; IPR000743; GH28.				
DR	Pfam; PF00295; Glyco_Hydro_28; 1.				

DR PROSITE; PS00502; POLYGALECTURONASE; 1.
 KW Hydrolyase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
 KM Amyloplase; Glycoprotein; Allergen
 FT SIGNAL 1 22
 FT PROPEP 23 45
 FT CHAIN 46 433
 FT PROPEP 434 514
 FT ACT SITE 278 278
 FT CARBOHYD 460 460
 FT CARBOHYD 472 472
 FT CONFLICT 5 5
 FT CONFLICT 12 12
 FT CONFLICT 34 35
 FT CONFLICT 37 37
 FT CONFLICT 88 88
 FT CONFLICT 98 98
 FT CONFLICT 451 451
 FT CONFLICT 454 454
 FT CONFLICT 504 504
 FT CONFLICT 507 507
 SQ SEQUENCE 514 AA; 56645 MW; 62461133FAD6302 CRC64;
 Query Match 100.0%; Score 81; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 VDGIIAAYONPASMK 15
 Db 120 VDGIIAAYONPASMK 134
 RESULT 2
 ACOC_ALCEU STANDARD; PRT; 373 AA.
 ID ACOC_ALCEU P27747;
 AC 01-AUG-1992 (Rel. 23, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dihydrolipoamide acetyltransferase component of acetoacetyl
 DE system (EC 2.3.1.12) (Acetoacetyl dehydrogenase E2 component)
 DE (Fast-migrating protein) (FMP).
 OS Alkaligenes eutrophus (Ralstonia eutropha).
 OS Bacteria; Proteobacteria; Beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=510;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-27.
 RC STRAIN=H16 / DSM 428 / ATCC 17699;
 RX MEDLINE=91266190; PubMed=2061286;
 RA Pfeleert H., Hein S., Krueger N., Zeh K., Schmidt B., Steinhuechel A.;
 RT "Identification and molecular characterization of the Alkaligenes
 RT eutrophus H16 acetoacetyl transferase gene involved in acetoacetyl
 RT J. Bacteriol. 173:4056-4071(1991).
 RL J. Bacteriol. 173:4056-4071(1991).
 CC -1- FUNCTION: DIHYDROLIPOAMIDE ACETYLTRANSFERASE INVOLVED IN
 CC ACETOIN CATABOLISM.
 CC CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-
 CC acetylhydrolipoamide.
 CC -1- COFACTOR: THE E2 COMPONENT CONTAINS ONE COVALENTLY-BOUND LIPOYL
 CC COFACTOR (PROBABLY).
 CC -1- PATHWAY: Acetoacetyl catabolism.
 CC -1- INDUCTION: BY GROWTH ON ACETOIN.
 CC -1- SIMILARITY: CONTAINS 1 LIPOYL-BINDING DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; H66060; AAA21950.1; -

DR FIR; D42462; D42462.
 DR HSSP; P20708; 1GHU.
 DR InterPro; IPR003089; AB hydrolyase.
 DR InterPro; IPR000073; Abhydrolyase.
 DR InterPro; IPR000089; Biotin lipoyl.
 DR InterPro; IPR003016; Lipoyl.
 DR InterPro; IPR000379; Ser esters site.
 DR Pfam; PF00364; biotin lipoyl; 1.
 DR Pfam; PF00561; abhydrolyase; 1.
 DR PRINTS; PR00111; ABHYDROLASE.
 DR PROSITE; PS00189; LIPOYL; 1.
 KW Transferase; Acyltransferase; lipoyl.
 FT INIT MET 0
 FT BINDING 49 49
 FT SEQUENCE 373 AA; 38904 MW; 9D5CE11B3E94892F CRC64;
 Query Match 55.6%; Score 45; DB 1; Length 373;
 Best Local Similarity 58.3%; Pred. No. 2;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 VDGIIAAYONP 12
 Db 92 IDGVAAVETPA 103
 RESULT 3
 BIME_EMENT STANDARD; PRT; 2073 AA.
 ID BIME_EMENT P24686;
 AC P24686;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Negative regulator of mitosis.
 DE BIME.
 OS Emericella nidulans (Aspergillus nidulans).
 OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90375468; PubMed=1697851;
 RA Engle D.B., Osman S.A., Osman A.H., Rosborough S., Xiang X.,
 RA Morris N.R.;
 RT "A negative regulator of mitosis in Aspergillus is a putative
 RT membrane-spanning protein."
 RT J. Biol. Chem. 265:16132-16137(1990).
 RL -1- FUNCTION: NEGATIVE REGULATOR OF MITOSIS IN E.NIDULANS. THIS
 RL PROTEIN IS PART OF A REGULATORY PATHWAY THAT INCLUDES THE NIMA
 RL PROTEIN KINASE. IT IS REQUIRED TO PREVENT PREMATURE ENTRY INTO
 RL MITOSIS AND PREVENT THEM FROM LEAVING MITOSIS.
 CC -1- SIMILARITY: TO MOUSE TSG24 AND YEAST YN1172W.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M59705; AA51478.1; -
 DR PIR; A37879; BWASBE.
 DR InterPro; IPR002015; APC proteasome.
 DR Pfam; PF01851; PC rep; 4.
 KW Transmembrane; Mitosis.
 FT DOMAIN 342 353
 FT TRANSMEM 1623 1643
 FT TRANSMEM 1665 1703
 FT TRANSMEM 1746 1764
 FT SEQUENCE 2073 AA; 229178 MW; 05E4B81EADDF51E4 CRC64;
 Query Match 53.1%; Score 43; DB 1; Length 2073;

Best Local Similarity 50.0%; Pred. No. 26;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 GIAYONPASW 14
DB 867 GVNASMNTASW 878

RESULT 4

CYSM_AQUAE STANDARD; PRT; 327 AA.
ID CYSM_AQUAE
AC 067507;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE (O-acetylserine (Thiol)-lyase) (CSASE)
DS CYSM OR AC1556.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificae; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shead M.J., Keller M., Anjay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus";
RL Nature 392:353-358(1998).
CC -1- CATALYTIC ACTIVITY: O3-acetyl-L-serine + H(2)S = L-cysteine +
CC acetate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: Cysteine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE CYSTEINE SYNTHASE/CYSTATHIONINE BETA-
CC SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AEO00745; AAC07459.1; -
CC DR HSSP; P12674; 10AS.
CC DR InterPro; IPR001926; B6_enzyme_beta.
CC DR InterPro; IPR001216; Cys_synthase.
CC DR Pfam; PF00291; PALP; 1.
CC DR TIGRPFAM; TIGR01136; cysKM; 1.
CC DR TIGRPFAM; TIGR01138; cysM; 1.
CC DR PROSITE; PS00901; CYS SYNTHASE; 1.
CC Lysase: Cysteine biosynthesis; Pyridoxal phosphate; Complete proteome.
CC FT BINDING 65 65
CC FT SEQUENCE 327 AA; 36155 MW; 54FC4338949FB40E CRC64;
CC -----
CC Query Match 50.6%; Score 41; DB 1; Length 327;
CC Best Local Similarity 75.0%; Pred. No. 9;
CC Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 YONPASWK 15
DB 167 YNNPASWK 174

RESULT 5
PGMU RAT STANDARD; PRT; 561 AA.
AC P38652;
DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phosphoglucosyltransferase (EC 5.4.2.2) (Glucose phosphomutase) (PGM).
GN PGM.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=Liver;
RX MEDLINE=94040821; PubMed=8224913;
RA Rivera A.A., Elton T.S., Dey N.B., Bouneille P., Marchase R.B.,
RT "Isolation and expression of a rat liver cDNA encoding
RT phosphoglucosyltransferase";
RL Gene 133:261-266(1993).

CC -1- FUNCTION: THIS ENZYME PARTICIPATES IN BOTH THE BREAKDOWN AND
CC SYNTHESIS OF GLUCOSE.
CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose 1-phosphate = alpha-D-glucose
CC 6-phosphate.
CC -1- COFACTOR: Magnesium (by similarity).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOHEXOSE MUTASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; L11694; AAA16862.1; -
DR PIR; JC2011; PMRT.
DR HSSP; P00949; 3PWG.
DR InterPro; IPR001485; PG/PWM_mutase.

DR Pfam; PF00408; PGM_PWM; 1.
DR Pfam; PF02878; PGM_PWM_1; 1.
DR Pfam; PF02879; PGM_PWM_1; 1.
DR Pfam; PF02880; PGM_PWM_1; 1.
DR PRINTS; PR00509; PGM_PWM.
DR PROSITE; PS00710; PGM_PWM; 1.
KW Isomerase; Phosphorylation; Magnesium.
FT INIT MET 0
FT ACT_SITE 116 116
FT SEQUENCE 561 AA; 61272 MW; 585A151D49170A66 CRC64;
FT BY SIMILARITY
FT FORMS THE PHOSPHOSERINE INTERMEDIATE
FT (BY SIMILARITY).

Query Match 50.6%; Score 41; DB 1; Length 561;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 IAYONPASWK 15
DB 349 IASYEPTGWK 359

RESULT 6
YRT3_CAEEL STANDARD; PRT; 544 AA.
ID YRT3_CAEEL
AC Q10046;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 60.5 kDa protein T07A5.3 in chromosome III.
GN T07A5.3.
OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-Bristol N2;
 RA Buck D.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO MAMMALIAN SODIUM/PHOSPHATE COTRANSPORTER 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z48055; CAA88134.1; -
 DR WormPeP; T07A5.3; CE01648.
 KM Hypothetical protein; Transport; Symport; Sodium transport;
 KM Transmembrane; Glycoprotein.
 FT TRANSMEM 50 70
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 128 148 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 194 214 POTENTIAL.
 FT TRANSMEM 221 241 POTENTIAL.
 FT TRANSMEM 286 306 POTENTIAL.
 FT TRANSMEM 324 344 POTENTIAL.
 FT TRANSMEM 361 381 POTENTIAL.
 FT TRANSMEM 385 405 POTENTIAL.
 FT TRANSMEM 417 437 POTENTIAL.
 FT TRANSMEM 451 471 POTENTIAL.
 SQ SEQUENCE 544 AA; 60543 MW; 4615B3C27D9F86CC CRC64;
 Query Match 50.0%; Score 40.5; DB 1; Length 544;
 Best Local Similarity 44.4%; Pred. No. 18;
 Matches 8; Conservative 2; Mismatches 5; Indels 3; Gaps 1;
 Oy 1 VDGIT--AAAYNPASWK 15
 Db 433 VGGVNTVTYQNDGK 450
 RESULT 7
 TRMD MYCTU STANDARD; PRT; 230 AA.
 ID TRMD_MYCTU
 AC Q10797;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE tRNA (Guanine-N(1)-)-methyltransferase (EC 2.1.1.31) (MIG-
 DE methyltransferase) (tRNA [Gm37] methyltransferase).
 GN TRMD OR RV2906C OR MT2974 OR MTCY274.37C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gae S., Barry C.E. III, Tekaia F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Kirog A., McLean J., Moule S., Murphy J.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.B., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Ueberback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SPECIFICALLY METHYLATES GUANOSINE-37 IN VARIOUS TRNAS
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
 CC homocysteine + tRNA containing N(1)-methylguanine.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO THE RNA METHYLTRANSFERASE TRMD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z74024; CAA98343.1; -
 DR TIGR; MT2974; -
 DR TubercuList; RV2906c; -
 DR InterPro; IPR002649; tRNA_mig_MT.
 DR Pfam; PF01746; tRNA_mig_MT; 1.
 DR ProDom; PD004978; tRNA_mig_MT; 1.
 DR TIGRFAMs; TIGR00088; trmd; 1.
 KM Transferase; Methyltransferase; tRNA processing; Complete proteome.
 SQ SEQUENCE 230 AA; 25166 MW; 5981C3D149EDDACS CRC64;

Oy 1 VDGIT--AAAYNPASWK 15
 Db 175 LDGLLEGPSYTRPASWR 191
 RESULT 8
 RE14 SCHPO STANDARD; PRT; 302 AA.
 ID RE14_SCHPO
 AC Q09150;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Meiotic recombination protein rec14.
 GN REC14 OR SPBC32F12.02.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97403302; PubMed=9258671;
 RA Evans D.H., Li Y.F., Fox M.E., Smith G.R.;
 RA "A WD repeat protein, Rec14, essential for meiotic recombination in
 RT Schizosaccharomyces pombe";
 RL Genetics 146:1253-1264 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Glymoures B.,
 RA Weltegen I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Mambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Gallberg F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovskiy G.V., Useyev D., Barrett B.G., Nurse P.,
 RL "The genome sequence of *Schizosaccharomyces pombe*."
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39144; AAB71433.1; -
 DR EMBL: AL023796; CAAL9363.1; -
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PRODOM: PD000018; WD40; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS00678; WD_REPEATS_1; FALSE_NEG.
 DR PROSITE: PS50082; WD_REPEATS_2; 3.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 DR KEGG: K01001; WD40; 7.
 DR Meiosis: Repeat; WD repeat.
 FT REPEAT 14 51 WD 1.
 FT REPEAT 57 96 WD 2.
 FT REPEAT 101 140 WD 3.
 FT REPEAT 142 184 WD 4.
 FT REPEAT 185 226 WD 5.
 FT REPEAT 227 266 WD 6.
 FT REPEAT 269 301 WD 7.
 SQ SEQUENCE 302 AA; 32936 MW; 7F650E98A7E1EB7 CRC64;
 Query Match 49.4%; Score 40; DB 1; Length 302;
 Best Local Similarity 53.3%; Pred. No. 12;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 VDGIIAAYONPASWK 15
 DB 123 VDGIIAAYONPASWK 137
 RESULT 9
 ID VB19_VACCD STANDARD; PRT; 351 AA.
 AC P23958;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Surface antigen S precursor (S antigen).
 GN B19R.
 OS Vaccinia virus (strain Dairen I).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 RN NCBI_TaxID=10250;
 RX

RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=90320131; PubMed=2196742;
 RX Ueda Y., Morikawa S., Matsura Y.,
 RT "Identification and nucleotide sequence of the gene encoding a
 RT surface antigen induced by vaccinia virus."
 RL Virology 177:588-594(1990).
 CC -1- FUNCTION: MAY BIND INTERLEUKIN-1 AND/OR INTERLEUKIN-6 AND PREVENT
 CC THESE CYTOKINES REACHING THEIR NATURAL RECEPTORS. IN CONSEQUENCE
 CC THE INFLAMMATORY RESPONSE WOULD BE DIMINISHED AND VIRUS
 CC REPLICATION ENHANCED.
 CC -1- SUBCELLULAR LOCATION: INDUCED ON THE SURFACE OF VACCINIA
 CC VIRUS-INFECTED CELLS.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: TO THE INTERLEUKIN-1 RECEPTOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D90076; BAA1416.1; -
 DR PIR: A35522; SANZVY.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003600; Ig_like.
 DR Pfam: PF00047; Ig; 2.
 DR SMART: SM00410; Ig_like; 1.
 DR KEGG: K01001; Ig; 2.
 DR Early protein.
 FT SIGNAL 1 19
 FT CHAIN 20 351
 FT DOMAIN 65 137
 FT DOMAIN 164 229
 FT DOMAIN 264 341
 FT DISULFID 73 129
 FT DISULFID 172 221
 FT DISULFID 272 333
 FT CARBOHYD 117 117
 FT CARBOHYD 182 182
 FT CARBOHYD 261 261
 FT CARBOHYD 269 269
 FT CARBOHYD 321 321
 SQ SEQUENCE 351 AA; 40701 MW; 20997CB67D39E7DB CRC64;
 Query Match 49.4%; Score 40; DB 1; Length 351;
 Best Local Similarity 28.6%; Pred. No. 15;
 Matches 4; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 QY 1 VDGIIAAYONPASWK 14
 DB 281 IDVLILEMENPSGM 294
 RESULT 10
 ID VB19_VACCV STANDARD; PRT; 351 AA.
 AC P25213;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Surface antigen S precursor (S antigen).
 GN B19R OR B18R.
 OS Vaccinia virus (strain WR).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OC NCBI_TaxID=10254;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=91259063; PubMed=2045793;
 RX

RA Smith G.L., Chan Y.S., Howard S.T.;
 RT "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
 RL the right inverted terminal repeat."
 RN J. Gen. Virol. 72:1349-1376(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91170931; PubMed=1826022;
 RA Smith G.L., Chan Y.S.;
 RT "Two vaccinia virus proteins structurally related to the
 RL interleukin-1 receptor and the immunoglobulin superfamily."
 CC J. Gen. Virol. 72:511-518(1991)
 CC -1- FUNCTION: MAY BIND INTERLEUKIN-1 AND/OR INTERLEUKIN-6 AND PREVENT
 CC THESE CYTOKINES REACHING THEIR NATURAL RECEPTORS. IN CONSEQUENCE
 CC THE INFLAMMATORY RESPONSE WOULD BE DIMINISHED AND VIRUS
 CC REPLICATION ENHANCED.
 CC -1- SUBCELLULAR LOCATION: INDUCED ON THE SURFACE OF VACCINIA
 CC VIRUS-INFECTED CELLS.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: TO THE INTERLEUKIN-1 RECEPTOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D11079; BAA01848.1; -
 CC EMBL: D01019; BAA00826.1; -
 CC EMBL: A19579; CAA01478.1; -
 CC PIR: B38472; SAVZWR.
 CC PIR: JQ1812; JQ1812.
 CC InterPro: IPR003006; IG_MHC.
 CC InterPro: IPR003600; IG_Like.
 CC Pfam: PF00047; Ig_2.
 CC SMART: SM00410; Ig_Like; 1.
 CC Immunoglobulin domain; Antigen; Glycoprotein; Repeat; Signal;
 KW Early protein.
 FT SIGNAL 1 19
 FT CHAIN 20 351
 FT DOMAIN 65 137
 FT DOMAIN 164 229
 FT DOMAIN 264 341
 FT DISULFID 73 129
 FT DISULFID 172 221
 FT DISULFID 272 333
 FT CARBOHYD 117 117
 FT CARBOHYD 182 182
 FT CARBOHYD 261 261
 FT CARBOHYD 269 269
 FT CARBOHYD 321 321
 FT CARBOHYD 40731 MW; 73BFD6EAE1EF73F91 CRC64;
 SQ SEQUENCE 351 AA;
 Query Match 49.4%; Score 40; DB 1; Length 351;
 Best Local Similarity 28.6%; Pred. No. 15;
 Matches 4; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 QY 1 VDGIIAAYONPASW 14
 DB 281 IDVLIVEMENPSGW 294
 RESULT 11
 VB19 VACCC STANDARD; PRT; 353 AA.
 ID P21077;
 AC 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Surface antigen S precursor (S antigen).
 GN B19R. "

OS Vaccinia virus (strain Copenhagen).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OC NCBI_TaxId=10249;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91021027; PubMed=2219722;
 RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
 RA Paolletti E.;
 RT "The complete DNA sequence of vaccinia virus."
 RL Virology 179:247-266(1990).
 CC [2]
 CC COMPLETE GENOME.
 CC Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
 CC Paolletti E.;
 CC "Appendix to 'The complete DNA sequence of vaccinia virus'";
 CC Virology 179:517-563(1990).
 CC -1- FUNCTION: MAY BIND INTERLEUKIN-1 AND/OR INTERLEUKIN-6 AND PREVENT
 CC THESE CYTOKINES REACHING THEIR NATURAL RECEPTORS. IN CONSEQUENCE
 CC THE INFLAMMATORY RESPONSE WOULD BE DIMINISHED AND VIRUS
 CC REPLICATION ENHANCED.
 CC -1- SUBCELLULAR LOCATION: INDUCED ON THE SURFACE OF VACCINIA
 CC VIRUS-INFECTED CELLS.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: TO THE INTERLEUKIN-1 RECEPTOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M35027; AAA48218.1; -
 CC PIR: I42527; SAVZVC.
 CC InterPro: IPR003006; IG_MHC.
 CC InterPro: IPR003600; IG_Like.
 CC Pfam: PF00047; Ig_2.
 CC SMART: SM00410; Ig_Like; 1.
 CC Immunoglobulin domain; Antigen; Glycoprotein; Repeat; Signal;
 KW Early protein.
 FT SIGNAL 1 21
 FT CHAIN 22 353
 FT DOMAIN 67 139
 FT DOMAIN 166 231
 FT DOMAIN 266 343
 FT DISULFID 75 131
 FT DISULFID 174 223
 FT DISULFID 274 335
 FT CARBOHYD 119 119
 FT CARBOHYD 184 184
 FT CARBOHYD 263 263
 FT CARBOHYD 271 271
 FT CARBOHYD 323 323
 FT CARBOHYD 40952 MW; 51BFB080CDE1A224 CRC64;
 SQ SEQUENCE 353 AA;
 Query Match 49.4%; Score 40; DB 1; Length 353;
 Best Local Similarity 28.6%; Pred. No. 15;
 Matches 4; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 QY 1 VDGIIAAYONPASW 14
 DB 283 IDVLIVEMENPSGW 296
 RESULT 12
 GLGA DEIRA STANDARD; PRT; 444 AA.
 ID GLGA DEIRA
 AC O9RWS1;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glycogen synthase (EC 2.4.1.21) (Starch [bacterial glycogen]
 GN synthase).
 GN GLGA OR DR0594.
 OS Deinococcus radiodurans.
 CC Bacteria; Thermus/deinococcus group; Deinococci; Deinococcales;
 CC Deinococcaceae; Deinococcus.
 CX NCBI_TaxID=1299;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffitt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Makarewicz J.J., Lam P., McDonald L., Ueberback T., Zalewski C.,
 RA Vamatheva K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RA "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1";
 RL Science 286:1571-1577(1999).
 CC -1- FUNCTION: Synthesizes alpha-1,4-glucan chains using ADP-glucose.
 CC -1- CATALYTIC ACTIVITY: ADP-glucose + (1,4)-alpha-D-glucosyl(N) =
 CC ADP + (1,4)-alpha-D-glucosyl(N+1).
 CC -1- PATHWAY: Glycogen biosynthesis; second step.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AE001917; AAF10170.1; ALT_INT.
 DR TIGR; DR0594;
 DR InterPro; IPR001296; Glycos_transf_1.
 DR Pfam; PF00534; Glycos_transf_1; 2.
 KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
 KM Complete proteome.
 FT BINDING 15 ADP-GLUCOSE (BY SIMILARITY).
 SQ SEQUENCE 444 AA; 48457 MW; 767605781A915302 CRC64;
 Query Match 49.4%; Score 40; DB 1; Length 444;
 Best Local Similarity 60.0%; Pred. No. 18;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 6 AAYQNPASWK 15
 DB 408 AAFQDPADQWQ 417
 RESULT 13
 PRO1 DICDI STANDARD; PRT; 126 AA.
 AC P26199;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Profilin I.
 GN PROA.
 OS Dictyostelium discoideum (Slime mold).
 CC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 CX NCBI_TaxID=44689;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=AX2;
 RX MEDLINE=92226170; PubMed=1725525;
 RA Haugwitz M., Noegel A.A., Rieger D., Lottspeich F., Schleicher M.;
 RA "Dictyostelium discoideum contains two profilin isoforms that differ

RT in structure and function";
 RL J. Cell Sci. 100:481-489(1991).
 CC -1- FUNCTION: BINDS TO ACTIN AND AFFECTS THE STRUCTURE OF THE
 CC CYTOSKELETON. AT HIGH CONCENTRATIONS, PROFILIN PREVENTS THE
 CC POLYMERIZATION OF ACTIN, WHEREAS IT ENHANCES IT AT LOW
 CC CONCENTRATIONS. BY BINDING TO PIP2, IT INHIBITS THE FORMATION OF
 CC IP3 AND DG.
 CC -1- SUBUNIT: OCCURS IN MANY KINDS OF CELLS AS A COMPLEX WITH MONOMERIC
 CC ACTIN IN A 1:1 RATIO.
 CC -1- SIMILARITY: BELONGS TO THE PROFILIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X61581; CAA43781.1; -.
 DR PIR; S18027; PADO1.
 DR HSSP; Q42449; LAOK.
 DR DictyDB; DD01035; PROA.
 DR InterPro; IPR002097; Profilin.
 DR Pfam; PF00235; Profilin; 1.
 DR SMART; SM00392; PROFILIN; 1.
 DR PROSITE; PS00414; PROFILIN; 1.
 KM Actin-binding; Cytoskeleton; Multigene family.
 SQ SEQUENCE 126 AA; 13064 MW; 0A92CAD75D0CE970 CRC64;
 Query Match 46.1%; Score 39; DB 1; Length 126;
 Best Local Similarity 63.6%; Pred. No. 7.8;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 DGIYAAYQNP 12
 DB 44 DGIALFQNP 54
 RESULT 14
 CARB_STRTH STANDARD; PRT; 299 AA.
 AC P13079;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE rRNA methyltransferase (EC 2.1.1.-) (Carbomycin-resistance protein).
 GN CARB.
 OS Streptomyces thermotolerans.
 OG Streptomyces thermotolerans.
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 CX NCBI_TaxID=80858;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11416;
 RX MEDLINE=87248111; PubMed=3036668;
 RA Epp J.K., Burgett S.G., Schoner B.E.;
 RA "Cloning and nucleotide sequence of a carbomycin-resistance gene from
 RT Streptomyces thermotolerans.";
 RL Gene 53:73-83(1987).
 CC -1- FUNCTION: PROBABLE RNA METHYLASE. CARB CONFERS RESISTANCE TO
 CC CARBOMYCIN AND SEVERAL OTHER MACROLIDES, LINCOMYCIN AND
 CC VERNAMYCIN B, BUT NOT TO ALL MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN
 CC B ANTIBIOTICS.
 CC -1- INDUCTION: THE EXPRESSION OF CARB IS INDUCIBLE BY CERTAIN
 CC MACROLIDE ANTIBIOTICS.
 CC -1- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M16503; AAC32026.1; -
DR PIR; A26512; A26512.
DR InterPro; IPR001737; RNA_A_dimeth.
DR InterPro; IPR000051; SAM_Bind.
DR Pfam; PF00398; RnaM; 1.
DR PROSITE; PS01131; RNA_A_DIMETH; 1.
KW Antibiotic resistance; Transferase; Methyltransferase; Plasmid.
SQ SEQUENCE 299 AA; 34138 MW; 63990AA894C044C6 CRC64;

Query Match 48.1%; Score 39; DB 1; Length 299;
Best Local Similarity 41.7%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 GI1AAYQNPASW 14
Db 278 GVVVAVVTPEQW 289

RESULT 15
GPD_A_TREPA STANDARD; PRT; 356 AA.

AC 083973;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94) (NAD(P)H-
dependent glycerol-3-phosphate dehydrogenase).
GN GSA OR TP1009.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.

OX NCBI_Taxid=160;
RN 11)
RP SEQUENCE FROM N.A.

RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren B., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;

RT "Complete genome sequence of Treponema pallidum, the syphilis
spirochete.";
RL Science 281:375-388 (1998).

CC -1- CATALYTIC ACTIVITY: Sn-glycerol 3-phosphate + NAD(P) (+) =
CC glycerone phosphate + NAD(P)H.

CC -1- PATHWAY: De novo phospholipid biosynthesis; glycerol-3 phosphate
CC formation.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT GLYCEROL-3-PHOSPHATE
CC DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE001269; AAC65960.1; -
CC TIGR; TP1009; -

DR InterPro; IPR001652; NAD_Gly3P_dh.

DR Pfam; PF01210; NAD_Gly3P_dh; 1.

DR PRINTS; PR00077; GPDHDEGNASE.

DR PRODom; PD001649; NAD_Gly3P_dh; 1.
DR PROSITE; PS00957; NAD_G3PDH; FALSE_NEG.

KW Phospholipid biosynthesis; Oxidoreductase; NAD; Complete proteome.
SQ SEQUENCE 356 AA; 37878 MW; 03967CB61767D915 CRC64;

Query Match 48.1%; Score 39; DB 1; Length 356;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VDGI1AAYQNP 11
Db 163 VTGLIAASQNP 173

Search completed: April 20, 2003, 13:07:38
Job time : 3.92105 secs

09akr0 clostridium
09tlv12 clostridium
09tlv8 lupula herp
08xyd6 anabaena sp
08xyd7 anabaena sp
08xyw3 anabaena sp
09tse93 caenothax
05tj33 oerikovia x
09etv4 blumeria gr
08u964 agrobacteri
08t096 yellow head
08xy9 anabaena sp
08t4r2 monkeypov
09k4f6 pisum sativ
08r2n8 camelpx vi
09asn3 oryza sativ
08r7b0 salmonella
08p63 salmonella
08xb68 escherichia
02t919 tetrahymena
09wy15 thermococ
09cxb7 dirosophila
055394 synechocyst
09tuf7 bean 58058
09cst1 bacillus ha
072754 compox viru
09fj3a vaccinia vi
09fj85 ectromella
09lepa ectromella

Query Match	55.6%;	Score 45;	DB 10;	Length 442;
Best Local Similarity	63.6%;	Pred. No. 17;		
Matches	7;	Conservative	3;	Mismatches 1;
				Indels 0;
				Gaps 0.

```

DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Gb|AAC24188.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneo T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20363039; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones."
RL DNA Res. 7:217-221(2000).
DR EMBL; AP000382; BAB02940.1; -.
DR InterPro; IPR003653; SUMO_protease.

```

DR Pfam; PF02902; Peptidase_C48; 1.
 SQ SEQUENCE 1119 AA; 125821 MW; CD3483A918B0895 CRC64;
 Query Match 55.6%; Score 45; DB 10; Length 1119;
 Best Local Similarity 63.6%; Pred. No. 51;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAYONP 11
 |||:|:|:|
 Db 226 VDGVAHNSNP 236

RESULT 6
 ID Q9ZPF3 PRELIMINARY; PRT; 1175 AA.
 AC Q9ZPF3;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE T3H13.8 protein.
 GN T3H13.8 OR A74G08880.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA WASHU;
 RT "The A. thaliana Genome Sequencing Project.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Drone K.; Nguyen C.;
 RT "The sequence of A. thaliana T3H13.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Waterston R.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A7128396; AAD17370.1;
 DR EMBL; AL161513; CAB78012.1;
 DR InterPro; IPR003653; SUMO_protease.
 DR Pfam; PF02902; Peptidase_C48; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1175 AA; 131492 MW; AD1924338690E242 CRC64;

Query Match 55.6%; Score 45; DB 10; Length 1175;
 Best Local Similarity 63.6%; Pred. No. 54;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAYONP 11
 |||:|:|:|
 Db 90 VDGVAHNSNP 100

RESULT 7
 ID Q9C6N9 PRELIMINARY; PRT; 1198 AA.
 AC Q9C6N9;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 133.6 kDa protein.
 GN F2815.2.
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A.; Ecker J.R.; Palm C.J.; Federspiel N.A.; Kaul S.;
 RA White O.; Alonso J.; Altati H.; Araujo R.; Bowman C.L.; Brooks S.Y.;
 RA Buehler E.; Chan A.; Chao Q.; Chen H.; Cheuk R.F.; Chin C.W.;
 RA Chung M.K.; Conn L.; Conway A.B.; Conway A.R.; Creasy T.H.; Dewar K.;
 RA Dunn P.; Egu P.; Feldblum T.V.; Feng J.-D.; Fong B.; Fujita C.Y.;
 RA Gill J.E.; Goldsmith A.D.; Haas B.; Hansen N.F.; Hughes B.; Hultzer L.;
 RA Hunter J.L.; Jenkins J.; Johnson-Hopson C.; Khan S.; Khaykin E.;
 RA Kim C.J.; Koo H.L.; Kremetska I.; Kurtz D.B.; Kwan A.; Lam B.;
 RA Langin-Hooper S.; Lee A.; Lee J.M.; Lenz C.A.; Li J.H.; Li Y.-P.;
 RA Lin X.; Liu S.X.; Liu Z.A.; Lueros J.S.; Maiti R.; Marziani A.;
 RA Miltseher J.; Miranda M.; Nguyen M.; Nierman W.C.; Osborne B.T.;
 RA Pai G.; Peterson J.; Pham P.K.; Rizzo M.; Rooney T.; Rowley D.;
 RA Sakano H.; Salzberg S.L.; Schwartz J.R.; Shinn P.; Southwick A.M.;
 RA Sun H.; Tallon L.J.; Tambunga G.; Toriumi M.J.; Town C.D.;
 RA Utterback T.; Van Aken S.; Vaysberg M.; Vysotskaya V.S.; Walker M.;
 RA Wu D.; Yu G.; Fraser C.M.; Venter J.C.; Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana.";
 RL Nature 408:816-820 (2000).
 DR EMBL; AC079280; AAG50582.1;
 DR InterPro; IPR003653; SUMO_protease.
 DR Pfam; PF02902; Peptidase_C48; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1198 AA; 133579 MW; D608AE577CE21339 CRC64;

Query Match 55.6%; Score 45; DB 10; Length 1198;
 Best Local Similarity 63.6%; Pred. No. 55;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDGIIAYONP 11
 |||:|:|:|
 Db 226 VDGVAHNSNP 236

RESULT 8
 ID Q9C607 PRELIMINARY; PRT; 1201 AA.
 AC Q9C607;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 134.0 kDa protein.
 GN F14G11.4.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A.; Ecker J.R.; Palm C.J.; Federspiel N.A.; Kaul S.;
 RA White O.; Alonso J.; Altati H.; Araujo R.; Bowman C.L.; Brooks S.Y.;
 RA Buehler E.; Chan A.; Chao Q.; Chen H.; Cheuk R.F.; Chin C.W.;
 RA Chung M.K.; Conn L.; Conway A.B.; Conway A.R.; Creasy T.H.; Dewar K.;
 RA Dunn P.; Egu P.; Feldblum T.V.; Feng J.-D.; Fong B.; Fujita C.Y.;
 RA Gill J.E.; Goldsmith A.D.; Haas B.; Hansen N.F.; Hughes B.; Hultzer L.;
 RA Hunter J.L.; Jenkins J.; Johnson-Hopson C.; Khan S.; Khaykin E.;
 RA Kim C.J.; Koo H.L.; Kremetska I.; Kurtz D.B.; Kwan A.; Lam B.;
 RA Langin-Hooper S.; Lee A.; Lee J.M.; Lenz C.A.; Li J.H.; Li Y.-P.;
 RA Lin X.; Liu S.X.; Liu Z.A.; Lueros J.S.; Maiti R.; Marziani A.;
 RA Miltseher J.; Miranda M.; Nguyen M.; Nierman W.C.; Osborne B.T.;
 RA Pai G.; Peterson J.; Pham P.K.; Rizzo M.; Rooney T.; Rowley D.;
 RA Sakano H.; Salzberg S.L.; Schwartz J.R.; Shinn P.; Southwick A.M.;

RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uteback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana"; 816-820 (2000).
 RL Nature 408:816-820 (2000).
 DR EMBL; AC084221; AAG50521.1; -;
 DR InterPro; IPR003653; SUMO protease.
 DR Pfam; PF02902; Peptidase_C48; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 1201 AA; 133995 MW; D10F9847289CFB7 CRC64;

Query Match 55.6%; Score 45; DB 10; Length 1201;
 Best Local Similarity 63.6%; Pred. No. 56;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VDGIIAAYQNP 11
 |||:|:|:
 DB 226 VDGVAHNSNP 236

RESULT 9
 O9ZR01 PRELIMINARY; PRT; 1285 AA.

AC O9ZR01. 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical 143.8 kDa protein.
 GN F4C21.23 OR AT4G03300.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Habermann K., de la Bastide M., Huang E.N., Gnoj L., Schurz K.,
 RA Preston R., Calma C., Martienssen R., Parnell L.D., Dedhia N.,
 RA McCombie W.R.;
 RT "Arabidopsis thaliana BAC F4C21 from chromosome IV near 17 cM";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
 RA Preston R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
 RA Shepherd M., Schurz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC005275; AAD14458.1; -;
 DR EMBL; AL161496; CAB77815.1; -;
 DR InterPro; IPR003653; SUMO protease.
 DR Pfam; PF02902; Peptidase_C48; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 1285 AA; 143833 MW; 7D1542C2E7E0F5DC CRC64;

Query Match 55.6%; Score 45; DB 10; Length 1285;
 Best Local Similarity 63.6%; Pred. No. 60;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VDGIIAAYQNP 11
 |||:|:|:
 DB 378 VDGVAHNSNP 388

RESULT 10
 O9ZU21 PRELIMINARY; PRT; 1305 AA.
 AC O9ZU21;
 O9ZU21;

DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE F5F19.8 protein.
 GN F5F19.8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Lenz C., Liu S.,
 RA Li J., Kremenetskaia I., Lueros J., Gonzalez A., Altafi H., Araujo R.,
 RA Buehler E., Conway A.B., Conn L., Dunn P., Hansen N., Huizar L.,
 RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
 RA Ecker J.R., Federle N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F5F19 sequence";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC006216; AAD12676.1; -;
 DR InterPro; IPR003653; SUMO protease.
 DR Pfam; PF02902; Peptidase_C48; 1.
 SQ SEQUENCE 1305 AA; 146080 MW; D3C036AEE2EFD6C CRC64;

Query Match 55.6%; Score 45; DB 10; Length 1305;
 Best Local Similarity 63.6%; Pred. No. 61;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VDGIIAAYQNP 11
 |||:|:|:
 DB 226 VDGVAHNSNP 236

RESULT 11
 O9M194 PRELIMINARY; PRT; 1314 AA.

AC O9M194. 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical 146.5 kDa protein.
 GN F7P3.20.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Vitale D., Ligouri R., Flores M., Argirion A., De Simone V.,
 RA Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,
 RA Salanoubat M.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL138663; CAB86693.1; -;
 DR InterPro; IPR003653; SUMO protease.
 DR Pfam; PF02902; Peptidase_C48; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 1314 AA; 146506 MW; 534C779C9A11BE CRC64;

Query Match 55.6%; Score 45; DB 10; Length 1314;
 Best Local Similarity 63.6%; Pred. No. 62;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VDGIIAAYQNP 11
 |||:~|:~|:
 DB 226 VDGVAHNSNP 236

RESULT 12

09LUG8
ID Q9LUG8 PRELIMINARY; PRT; 1444 AA.
AC Q9LUG8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similarity to mutator-like transposase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB025605; BAA98062.1; -
DR InterPro; IPR003653; SUMO_protease.
DR Pfam; PF02902; Peptidase_C48; 1.
SQ SEQUENCE 1444 AA; 161024 MW; AF0CB23AAAC4EB8B CRC64;
Query Match 55.6%; Score 45; DB 10; Length 1444;
Best Local Similarity 63.6%; Pred. No. 69;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
CY 1 VDGIIAYONP 11
Db 226 VDGVAASHNP 236
RESULT 13
Q9SFY0
ID Q9SFY0 PRELIMINARY; PRT; 1444 AA.
AC Q9SFY0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE T22C5.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shiao P., Alafati H., Bel O., Chin C., Chou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharly N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Tourani M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RT "Genomic sequence for Arabidopsis thaliana BAC T22C5 from chromosome
1."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC012375; AAF24952.1; -
DR InterPro; IPR003653; SUMO_protease.
DR Pfam; PF02902; Peptidase_C48; 1.
SQ SEQUENCE 1444 AA; 161065 MW; C0563453E81C6498 CRC64;
Query Match 55.6%; Score 45; DB 10; Length 1444;
Best Local Similarity 63.6%; Pred. No. 69;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
CY 1 VDGIIAYONP 11
Db 226 VDGVAASHNP 236
RESULT 14

080975
ID 080975 PRELIMINARY; PRT; 1756 AA.
AC 080975;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE At2g14770 protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bentto M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.-U., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana."
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004705; AAC24188.1; -
DR InterPro; IPR003653; SUMO_protease.
DR Pfam; PF02902; Peptidase_C48; 1.
SQ SEQUENCE 1756 AA; 197368 MW; 68FC3B5FCC282775 CRC64;
Query Match 55.6%; Score 45; DB 10; Length 1756;
Best Local Similarity 63.6%; Pred. No. 87;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
CY 1 VDGIIAYONP 11
Db 226 VDGVAASHNP 236
RESULT 15
061266
ID 061266 PRELIMINARY; PRT; 521 AA.
AC 061266;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Matrix metalloproteinase.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RX MEDLINE=98241501; PubMed=9573338;
RA Wada K., Sato H., Kinoh H., Kajita M., Yamamoto H., Seiki M.;
RT "Cloning of three Caenorhabditis elegans genes potentially encoding
novel matrix metalloproteinases."
RL Gene 211:57-62 (1998)
DR EMBL; AB007817; BAA28353.1; -
DR HSRF; P03956; 1CGU.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Macrixin.
DR InterPro; IPR000130; Zn_MTPeptide.
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF00413; Peptidase_M10; 1.

DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZINC; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 521 AA; 60700 MW; F3936A1C732C6FE2 CRC64;

Query Match 54.3%; Score 44; DB 5; Length 521;
Best Local Similarity 40.0%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 VDGIIAAYQNPASWK 15
:|::|||
Db 229 IDSVWFAYDTPRKWK 243

Search completed: April 20, 2003, 13:13:10
Job time : 15.6711 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds
(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524d-100

Perfect score: 82

Sequence: 1 NNRIWLOPAKLTGFT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.73.*
2: PIR1.*
3: PIR2.*
4: PIR3.*
5: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	82	100.0	514	2	S48730	Cry j II protein -
2	82	100.0	514	2	JC2498	second major aller
3	58	70.7	514	2	JC7100	polygalacturonase
4	52	63.4	507	2	JC7366	Jun a 2 protein -
5	47	57.3	1161	2	H86368	protein F28C11.9 (
6	43	52.4	460	2	H96728	probable polygalac
7	42	51.2	266	2	A12283	hypothetical prote
8	42	51.2	492	2	C96521	protein F21D18.18
9	42	51.2	540	2	B96631	probable polygalac
10	42	51.2	853	2	S59315	hypothetical prote
11	41	50.0	63	2	D34514	MHC class II histo
12	41	50.0	235	2	D65071	hypothetical prote
13	41	50.0	235	2	B91097	hypothetical prote
14	41	50.0	235	2	A85943	hypothetical prote
15	41	50.0	452	1	D69810	phosphotransferase
16	40	48.8	246	2	F72464	hypothetical prote
17	40	48.8	756	2	S74742	exopolysaccharide
18	39	47.6	217	2	AB3432	thiol:disulfide in
19	39	47.6	225	2	T39083	conserved hypotnet
20	39	47.6	475	2	G82371	Flxg-related prote
21	39	47.6	677	2	T26514	hypothetical prote
22	39	47.6	880	2	G90163	DNA-directed RNA p
23	39	47.6	1037	2	B65022	acridflavin resist
24	39	47.6	1037	2	D91045	antimycoside eff
25	39	47.6	1037	2	G85889	hypothetical prote
26	39	47.6	1037	2	AB0816	probable efflux pu
27	38	46.3	328	2	T24705	hypothetical prote
28	38	46.3	328	2	G64610	probable transcrip
29	38	46.3	328	2	G71504	probable transcrip

30	38	46.3	420	2	C71407	hypothetical prote
31	38	46.3	479	2	UQ0781	sucrose uptake pro
32	38	46.3	487	1	QOECRS	ylgE protein - Esc
33	38	46.3	549	2	F82415	conserved hypotnet
34	38	46.3	624	2	S50650	hypothetical prote
35	38	46.3	790	2	T50337	homolog to drosoph
36	38	46.3	1044	2	B72338	galactosyltransfer
37	37.5	45.7	493	2	G69468	carbohydrate kinas
38	37.5	45.7	525	2	H81386	ABC transporter AT
39	37.5	45.7	1035	2	AD3203	autotransporter pr
40	37	45.1	131	2	AE0126	conserved hypotnet
41	37	45.1	174	2	B72748	hypothetical prote
42	37	45.1	184	1	G64018	hypothetical prote
43	37	45.1	219	2	F71363	probable amino aci
44	37	45.1	241	2	T27636	hypothetical prote
45	37	45.1	299	2	F97277	probable glycosylt

ALIGNMENTS

RESULT 1

S48730 Cry j II protein - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 07-May-1995 #sequence_rev1501 21-Jul-1995 #text_change 28-May-1999

C/Accession: S48730

R/Namba, M.; Kurose, M.; Torijoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.;

FEBS Lett. 353, 124-128, 1994

A/Title: Molecular cloning of the second major allergen, Cry j II, from Japanese ceda

A/Reference number: S48730; MWID:95010777; PMID:7926035

A/Accession: S48730

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-514 <KOM>

A/Cross-references: GB:D37765; NID:9577695; PIDN:BAA07021.1; PID:dl007598; PID:957769

A/Cross-references: GB:D37765; NID:9577695; PIDN:BAA07021.1; PID:dl007598; PID:957769

Query Match Best Local Similarity 100.0%; Score 82; DB 2; Length 514;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNRIWLOPAKLTGFT 15

Db 135 NNRIWLOPAKLTGFT 149

RESULT 2

JC2498 second major allergen Cry j II precursor - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 16-Mar-1995 #sequence_rev1501 26-May-1995 #text_change 21-Jul-2000

C/Accession: JC2498; PC2346; A60147

R/Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994

A/Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanes

A/Reference number: JC2498; MWID:94271186; PMID:8002972

A/Accession: JC2498

A/Molecule type: mRNA

A/Residues: 1-514 <KOM>

A/Cross-references: DDBJ:D29772; NID:9506857; PIDN:BAA06172.1; PID:9506858

A/Accession: PC2346

A/Molecule type: protein

A/Residues: 52-61 <KO2>

R/Sakaguchi, M.; Inouye, S.; Tanhai, M.; Ando, S.; Usui, M.; Matuhasi, T.

Allergy 45, 309-312, 1990

A/Title: Identification of the second major allergen of Japanese cedar pollen.

A/Reference number: A60147; MWID:90342988; PMID:2382797

A/Accession: A60147

A/Molecule type: protein

A/Residues: 55-64 <SAK>

C/Keywords: glycoprotein; pollen

F1-54/Domains: signal sequence #status predicted <SIG>

F155-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 82; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNRIMLOPAKLTGFT 15
|||
Db 135 NNRIMLOPAKLTGFT 149

RESULT 3

UC7100 polygalacturonase Cha o 2 - Japanese cypress
C:Species: Chamaecyparis obtusa (Japanese cypress)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C:Accession: JG7100; PMID:99417540; PMID:10486372
R:Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.
Biochem. Biophys. Res. Commun. 263, 166-171, 1999
A:Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all
A:Reference number: JG7100; PMID:99417540; PMID:10486372
A:Accession: JG7100
A:Molecule type: mRNA
A:Residues: 1-514 <MOR>
A:Accession: PC7026
A:Molecule type: protein
A:Residues: 51-62 <MO2>

Query Match 70.7%; Score 58; DB 2; Length 514;
Best Local Similarity 71.4%; Pred. No. 0.022;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNRIMLOPAKLTGFT 14
|||
Db 135 NNRIMLOPAKLTGFT 148

RESULT 4

UC7366 Jun a 2 protein - mountain cedar
C:Species: Juniperus ashei (mountain cedar)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JG7366; PC7093
R:Yokoyama, M.; Miyahara, M.; Shintzu, K.; Kino, K.; Tsunoo, H.
Biochem. Biophys. Res. Commun. 275, 195-202, 2000
A:Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all
A:Reference number: JG7366
A:Accession: JG7366
A:Molecule type: mRNA
A:Residues: 1-507 <YOK>
A:Cross-references: GB:AJ046653
A:Accession: PC7093
A:Molecule type: protein
A:Residues: 55-63 <Y02>
C:Comment: This protein, a second major allergen of mountain cedar pollen, which is inv
C:Keywords: glycoprotein; pollen

Query Match 63.4%; Score 52; DB 2; Length 507;
Best Local Similarity 57.1%; Pred. No. 0.26;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 NNRIMLOPAKLTGFT 14
|||
Db 136 NNRIMLOPAKLTGFT 149

RESULT 5

B86368 protein F28C11.9 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B86368

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, I.
Nansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marzida,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; PMID:21016719; PMID:11130712
A:Accession: B86368
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 11161 <STO>
A:Cross-references: GB:AE005173; NID:98778576; PIDN:AAF79584.1; GSPDB:GN00141
C:Genetics:
A:Gene: F28C11.9
A:Map position: 1

Query Match 57.3%; Score 47; DB 2; Length 1161;
Best Local Similarity 72.7%; Pred. No. 4.9;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RIMLOPAKLTG 13
|||
Db 824 RIMLOPAKLTG 834

RESULT 6

H96728 probable polygalacturonase F24J13.7 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96728
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marzida,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; PMID:21016719; PMID:11130712
A:Accession: H96728
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-468 <STO>
A:Cross-references: GB:AE005173; NID:96175142; PIDN:AAF04869.1; GSPDB:GN00141
C:Genetics:
A:Gene: F24J13.7
A:Map position: 1

Query Match 52.4%; Score 43; DB 2; Length 468;
Best Local Similarity 72.7%; Pred. No. 9.7;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RIMLOPAKLTG 13
|||
Db 147 RIMLOPAKLTG 157

RESULT 7

AI2283 hypothetical protein alr3824 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AI2283
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriy
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
 A:Reference number: AB1807, MUID:21595285, PMID:11759840
 A:Accession: AI2283
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-266 <KUR>
 A:Cross-references: GB:BA000019, PIDN:BA875523.1, PID:gi7132958, GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr3824

MHC class II histocompatibility antigen, 1 - rhesus macaque (fragment)

C/Species: Macaca mulatta (rhesus macaque)

C/Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 12-Jul-1996

C/Accession: D34514

R/Gyllenstein, U.B.; Erlich, H.A.

A/Title: Ancient roots for polymorphism at the HLA-DQA1 locus in primates.

A/Reference number: A34512; MUID:9009387; PMID:2513578

A/Accession: D34514

A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A/Molecule type: DNA

A/Residues: 1-63 <GVL>

C/Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 50.0%; Score 41; DB 2; Length 63;

Best Local Similarity 54.5%; Pred. No. 2.6;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 WLOPAKLTGFT 14

Db 30 LWFSKRGGF 40

RESULT 12

hypothetical protein b2876 - Escherichia coli (strain K-12)

C/Species: Escherichia coli

C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C/Accession: D65071

R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

A/Title: The complete genome sequence of Escherichia coli K-12.

A/Reference number: A64720; MUID:97426617; PMID:9278503

A/Accession: D65071

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-235 <BLAT>

A/Cross-references: GB:AE00371; GB:U00096; NID:q1789239; PIDN:AACT5914.1; PID:91789241;

A/Experimental source: strain K-12, substrain MG1655

C/Superfamily: Escherichia coli hypothetical protein b2876

Query Match 50.0%; Score 41; DB 2; Length 235;

Best Local Similarity 63.6%; Pred. No. 11;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 WLOPAKLTGFT 15

Db 147 WSOFAADITGLT 157

RESULT 13

hypothetical protein ECs3749 [imported] - Escherichia coli (strain O157:H7, substrain RI

C/Species: Escherichia coli

C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C/Accession: E91097

R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.

A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A/Reference number: A99629; MUID:21156231; PMID:11258796

A/Accession: E91097

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-235 <HAY>

A/Cross-references: GB:BA000007; PIDN:BA037172.1; PID:G13363221; GSPDB:GN00154

A/Experimental source: strain O157:H7, substrain RIMD 0509952

C/Genetics:

A/Genes: ECs3749

C/Superfamily: Escherichia coli hypothetical protein b2876

Query Match 50.0%; Score 41; DB 2; Length 235;

Qy 5 WLOPAKLTGFT 15

Db 147 WSOFAADITGLT 157

RESULT 14

hypothetical protein Z4215 [imported] - Escherichia coli (strain O157:H7, substrain ET

C/Species: Escherichia coli

C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C/Accession: A85943

R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayh

A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A/Reference number: A85480; MUID:21074935; PMID:11206551

A/Accession: A85943

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-235 <STO>

A/Cross-references: GB:AE005174; NID:G12517397; PIDN:AAGS8005.1; GSPDB:GN00145; UWGP:Z

A/Experimental source: strain O157:H7, substrain EDL933

C/Genetics:

A/Genes: Z4215

C/Superfamily: Escherichia coli hypothetical protein b2876

Query Match 50.0%; Score 41; DB 2; Length 235;

Qy 5 WLOPAKLTGFT 15

Db 147 WSOFAADITGLT 157

RESULT 15

phosphotransferase system enzyme II (EC 2.7.1.69) factor II homolog yJlf - Bacillus su

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C/Accession: D69810

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bery

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A/Reference number: A69580; MUID:98044035; PMID:9384377

A/Accession: D69810

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-452 <KUN>

A/Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12599.1; PID:G263305

A/Experimental source: strain 168

C/Genetics:

A/Genes: yJlf

C/Function:

A/Description: mediates transport of glucose across the cytoplasmic membrane concomit

C/Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosph

C/Keywords: phosphoprotein; phosphotransferase; sugar transport system; transmembrane

Query Match 50.0%; Score 41; DB 2; Length 235;

Qy 5 WLOPAKLTGFT 15

Db 147 WSOFAADITGLT 157

RESULT 16

phosphotransferase system enzyme II (EC 2.7.1.69) factor II homolog yJlf - Bacillus su

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C/Accession: D69810

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bery

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A/Reference number: A69580; MUID:98044035; PMID:9384377

A/Accession: D69810

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-452 <KUN>

A/Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12599.1; PID:G263305

A/Experimental source: strain 168

C/Genetics:

A/Genes: yJlf

C/Function:

A/Description: mediates transport of glucose across the cytoplasmic membrane concomit

C/Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosph

C/Keywords: phosphoprotein; phosphotransferase; sugar transport system; transmembrane

Query Match 50.0%; Score 41; DB 2; Length 235;

Qy 5 WLOPAKLTGFT 15

Db 147 WSOFAADITGLT 157

RESULT 17

phosphotransferase system enzyme II (EC 2.7.1.69) factor II homolog yJlf - Bacillus su

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C/Accession: D69810

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bery

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A/Reference number: A69580; MUID:98044035; PMID:9384377

A/Accession: D69810

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-452 <KUN>

A/Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12599.1; PID:G263305

A/Experimental source: strain 168

C/Genetics:

A/Genes: yJlf

C/Function:

A/Description: mediates transport of glucose across the cytoplasmic membrane concomit

C/Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosph

C/Keywords: phosphoprotein; phosphotransferase; sugar transport system; transmembrane

Query Match 50.0%; Score 41; DB 2; Length 235;

Qy 5 WLOPAKLTGFT 15

Db 147 WSOFAADITGLT 157

RESULT 18

phosphotransferase system enzyme II (EC 2.7.1.69) factor II homolog yJlf - Bacillus su

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C/Accession: D69810

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bery

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A/Reference number: A69580; MUID:98044035; PMID:9384377

A/Accession: D69810

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-452 <KUN>

A/Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12599.1; PID:G263305

A/Experimental source: strain 168

C/Genetics:

A/Genes: yJlf

C/Function:

A/Description: mediates transport of glucose across the cytoplasmic membrane concomit

C/Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosph

C/Keywords: phosphoprotein; phosphotransferase; sugar transport system; transmembrane

Query Match 50.0%; Score 41; DB 2; Length 235;

Qy 5 WLOPAKLTGFT 15

Db 147 WSOFAADITGLT 157

RESULT 19

phosphotransferase system enzyme II (EC 2.7.1.69) factor II homolog yJlf - Bacillus su

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C/Accession: D69810

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bery

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A/Reference number: A69580; MUID:98044035; PMID:9384377

A/Accession: D69810

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-452 <KUN>

A/Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12599.1; PID:G263305

A/Experimental source: strain 168

C/Genetics:

A/Genes: yJlf

C/Function:

A/Description: mediates transport of glucose across the cytoplasmic membrane concomit

C/Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosph

C/Keywords: phosphoprotein; phosphotransferase; sugar transport system; transmembrane

Query Match 50.0%; Score 41; DB 2; Length 235;

Qy 5 WLOPAKLTGFT 15

Db 147 WSOFAADITGLT 157

RESULT 20

phosphotransferase system enzyme II (EC 2.7.1.69) factor II homolog yJlf - Bacillus su

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C/Accession: D69810

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bery

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A/Reference number: A69580; MUID:98044035; PMID:9384377

A/Accession: D69810

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-452 <KUN>

A/Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12599.1; PID:G263305

A/Experimental source: strain 168

C/Genetics:

A/Genes: yJlf

C/Function:

A/Description: mediates transport of glucose across the cytoplasmic membrane concomit

C/Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosph

C/Keywords: phosphoprotein; phosphotransferase; sugar transport system; transmembrane

Query Match 50.0%; Score 41; DB 2; Length 235;

Qy 5 WLOPAKLTGFT 15

Db 147 WSOFAADITGLT 157

RESULT 21

phosphotransferase system enzyme II (EC 2.7.1.69) factor II homolog yJlf - Bacillus su

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C/Accession: D69810

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bery

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A/Reference number: A69580; MUID:98044035; PMID:9384377

A/Accession: D69810

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-452 <KUN>

A/Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12599.1; PID:G263305

A/Experimental source: strain 168

C/Genetics:

A/Genes: yJlf

C/Function:

A/Description: mediates transport of glucose across the cytoplasmic membrane concomit

C/Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosph

C/Keywords: phosphoprotein; phosphotransferase; sugar transport system; transmembrane

Query Match 50.0%; Score 41; DB 2; Length 235;

Qy 5 WLOPAKLTGFT 15

F:1452/Domain: phosphotransferase system glucose-specific enzyme II, factor II homology
 F:12-28/Domain: transmembrane #status predicted <TM01>
 F:46-62/Domain: transmembrane #status predicted <TM02>
 F:94-110/Domain: transmembrane #status predicted <TM03>
 F:133-149/Domain: transmembrane #status predicted <TM04>
 F:253-269/Domain: transmembrane #status predicted <TM05>
 F:286-302/Domain: transmembrane #status predicted <TM06>
 F:329-345/Domain: transmembrane #status predicted <TM07>
 F:419-436/Domain: transmembrane #status predicted <TM08>
 F:397/Active site: Cys (phosphocysteine intermediate) #status predicted

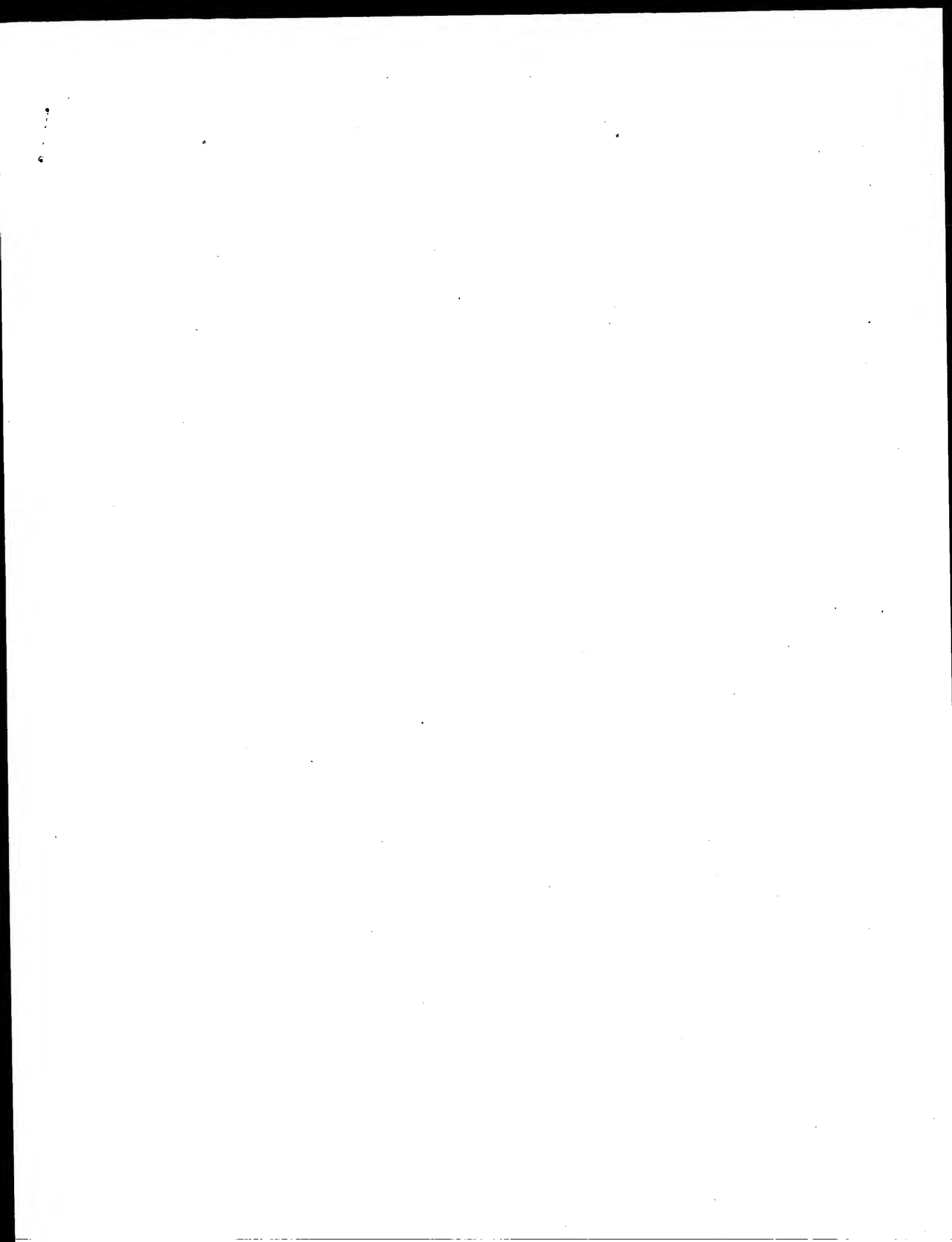
Query Match

Best Local Similarity 50.0%; Score 41; DB 1; Length 452;
 Pred. No. 21;

Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 NNRIMLPALITGFT 15
 Db 192 NNIFWFOFGEGYNGVT 206

Search completed: April 20, 2003, 13:15:51
 Job time : 8.07895 secs



GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524D-100

Perfect score: 82

Sequence: 1 NNR1WLOPAKLTGFT 15

Scoring table: BLOSUM62

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	514	1	MPA2_CRYVA
2	41	50.0	256	1	YOEC_ECOLI
3	39	47.6	880	1	RPAL_SULSO
4	39	47.6	1037	1	ACRD_ECOLI
5	38	46.3	479	1	P24177 escherichia
6	38	46.3	487	1	P22825 vibrio algi
7	38	46.3	624	1	PTDT_ECOLI
8	37	45.1	172	1	YEX7_YEAST
9	37	45.1	181	1	NADM_AERPE
10	37	45.1	310	1	YGAT_HAETN
11	37	45.1	359	1	PPP5_PSRSE
12	37	45.1	361	1	OTC_EMENT
13	37	45.1	370	1	OTC_ASPTE
14	37	45.1	409	1	OTC_ASPNG
15	37	45.1	490	1	HMDH_SULSO
16	37	45.1	778	1	CALR_HUMAN
17	37	45.1	982	1	EXP7_STRPN
18	37	45.1	983	1	ENV_VILV
19	37	45.1	989	1	ENV_VILV
20	37	45.1	990	1	ENV_VILV
21	37	45.1	991	1	ENV_VILV
22	37	45.1	1353	1	CYA9_HUMAN
23	37	45.1	1353	1	CYA9_MOUSE
24	36.5	44.5	372	1	PROB_THERN
25	36	43.9	122	1	SAA3_RABIT
26	36	43.9	178	1	RLIX_CASSA
27	36	43.9	178	1	RLIX_ORYSA
28	36	43.9	178	1	RLIX_ORYSA
29	36	43.9	255	1	UNG_HSVR2
30	36	43.9	315	1	CPP2_ENTHI
31	36	43.9	322	1	GM2_SCHPO
32	36	43.9	354	1	YHL4_YEAST
33	36	43.9	368	1	HIS8_BUCAI

ALIGNMENTS

RESULT 1	MPA2_CRYVA	STANDARD:	PRT:	514 AA.
ID	MPA2_CRYVA			
AC	P43212;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DE	Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)			
DE	(Major pollen allergen Cry j 2) (Cry j II).			
OS	Cryptomeria japonica (Japanese cedar).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.			
OX	NCBI_TaxId=3369;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=Pollen;			
RX	MEDLINE=9501077; PubMed=7926035;			
RA	Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,			
RT	Usui M., Kurimoto M.;			
RT	"Molecular cloning of the second major allergen, Cry j II, from			
RT	Japanese cedar pollen.";			
RL	FEBS Lett. 353:124-128 (1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pollen;			
RX	MEDLINE=94271186; PubMed=8002972;			
RA	Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.;			
RT	"CDNA cloning and expression of Cry j II the second major allergen of			
RT	Japanese cedar pollen.";			
RL	Biochem. Biophys. Res. Commun. 201:1021-1028(1994).			
RN	[3]			
RP	SEQUENCE OF 55-64.			
RC	Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matubasi T.;			
RT	"Identification of the second major allergen of Japanese cedar			
RT	pollen.";			
RL	Allergy 45:309-312(1990).			
CC	-I- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-			
CC	galactosiduronic linkages in pectate and other galacturonans.			
CC	-I- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).			
CC	-I- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES			
CC	(POLYGALACTONASES).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; D37765; BAA07021.1; -			
DR	EMBL; D29772; BAA06172.1; -			
DR	HSRP; P26509; IHR.			
DR	InterPro: IPR000743; GH28.			
DR	Pfam: PF00295; Glyco_hydro_28; 1.			

34	36	43.9	369	1	DHDO_DICDI	P07670 dictyosteli
35	36	43.9	478	1	CALR_GAYPO	O08893 cavia porce
36	36	43.9	522	1	FLUD_BACHD	O9660 bacillus ha
37	36	43.9	678	1	YHQ_ECOLI	P32138 escherichia
38	36	43.9	883	1	YHL6_YEAST	P38781 saccharomyc
39	36	43.9	974	1	SIN4_YEAST	P33259 saccharomyc
40	36	43.9	988	1	TERT_SCHPO	O13339 schizosacch
41	36	43.9	1046	1	MEXB_PSRSE	P53002 pseudomonas
42	36	43.9	1305	1	CYA9_XENLA	P06999 xenopus lae
43	36	43.9	1310	1	ICP4_VZVD	P09310 varicella-z
44	36	43.9	1334	1	CYA9_CHICK	O96996 gallus gall
45	36	43.9	1396	1	VLTG_BPTS	P13390 bacteriophag

DR PROSITE; PS00502; POLY GALACTURONASE; 1.
 KW Hydrolase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
 KW Amyloplast; Glycoprotein; Allergen.
 FT SIGNAL 1 22
 FT PROPEP 23 45
 FT CHAIN 46 433
 FT PROPEP 434 514
 FT ACT SITE 278 278
 FT CARBOHYD 460 460
 FT CARBOHYD 472 472
 FT CONFLICT 5 5
 FT CONFLICT 12 12
 FT CONFLICT 34 35
 FT CONFLICT 37 37
 FT CONFLICT 88 88
 FT CONFLICT 98 98
 FT CONFLICT 451 451
 FT CONFLICT 454 454
 FT CONFLICT 504 504
 FT CONFLICT 507 507
 SQ SEQUENCE 514 AA; 56645 MW; 6246113FAD6302 CRC64;

Query Match 100.0%; Score 82; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 9.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NNRWLOFAKLTGFT 15
 DB 135 NNRWLOFAKLTGFT 149

RESULT 2
 YOEBC_ECOLI STANDARD; PRT; 256 AA.
 AC Q46809;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein Yec.
 GN YOEBC OR B2876.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RA MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U28375; AAA83057.1; ALT_INIT.
 DR EMBL; AB000371; AAC75914.1; ALT_INIT.
 DR EcoGene; EG31059; YecC.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 256 AA; 28103 MW; F9E5C04EB26ED688 CRC64;

Query Match 50.0%; Score 41; DB 1; Length 256;
 Best Local Similarity 63.6%; Pred. No. 6.3;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 WLOFAKLTGFT 15
 DB 168 WSGFADITGLT 178

RESULT 3
 RPAI_SULSO STANDARD; PRT; 880 AA.
 AC Q980R2;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA-directed RNA polymerase subunit A' (EC 2.7.7.6).
 GN RPOA1 OR SSO0225.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RA MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moers A., Erasmo G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragam M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -1- COFACTOR: ZINC (By similarity).
 CC -1- SIMILARITY: THE COMBINED A'+A" SUBUNITS CORRESPOND TO THE A
 CC SUBUNITS OF EUKARYOTIC RNA POLYMERASES I, II AND III AND TO THE
 CC EUBACTERIAL BETA' SUBUNIT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB006559; AAK40566.1; -
 DR InterPro; IPR000722; RNA_pol_A.
 DR Pfam; PF00623; RNA_pol_A; 1.
 KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
 KW Zinc-finger; Complete proteome.
 FT ZN FING 58 101
 FT SEQUENCE 880 AA; 99674 MW; 5C8AE20FE71DE25 CRC64;

Query Match 47.6%; Score 39; DB 1; Length 880;
 Best Local Similarity 46.2%; Pred. No. 52;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 3 R1WLOFAKLTGFT 15
 DB 633 RVIRFVLEGT 645

RESULT 4
 ACRD_ECOLI STANDARD; PRT; 1037 AA.
 ID ACRD_ECOLI
 AC P24177; P77178; Q46715; P76971;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DR	EMBL; X57403; CAA00663.1; -.
DR	PfR; CA2959; CA2959.
DR	PfR; S26987; S26997.
DR	EcoGene; EGI0014; acid.
DR	InterPro; IPR001036; Aciflavin_res.
DR	InterPro; IPR004764; HAEI.
DR	Pfam; PF00873; ACR_tran; 1.
DR	PRINTS; PR00702; ACRIPLAVINRP.
DR	TIGRFAMs; TIGR00915; 2A0602; 1.
KM	Transmembrane; Inner membrane; Transport; Complete proteome.
FT	DOMAIN 1
FT	TRANSMEM 10
FT	DOMAIN 28
FT	TRANSMEM 29
FT	TRANSMEM 339
FT	TRANSMEM 340
FT	TRANSMEM 359
FT	TRANSMEM 360
FT	TRANSMEM 365
FT	TRANSMEM 366
FT	TRANSMEM 385
FT	TRANSMEM 386
FT	TRANSMEM 391
FT	TRANSMEM 392
FT	TRANSMEM 413
FT	DOMAIN 414
FT	TRANSMEM 442
FT	TRANSMEM 443
FT	TRANSMEM 460
FT	TRANSMEM 461
FT	TRANSMEM 473
FT	TRANSMEM 494
FT	TRANSMEM 497
FT	TRANSMEM 537
FT	TRANSMEM 538
FT	TRANSMEM 557
FT	TRANSMEM 871
FT	TRANSMEM 871
FT	TRANSMEM 890
FT	TRANSMEM 891
FT	TRANSMEM 896
FT	TRANSMEM 916
FT	TRANSMEM 917
FT	TRANSMEM 922
FT	TRANSMEM 923
FT	TRANSMEM 944
FT	TRANSMEM 945
FT	TRANSMEM 971
FT	TRANSMEM 992
FT	TRANSMEM 990
FT	TRANSMEM 991
FT	TRANSMEM 1003
FT	TRANSMEM 1004
FT	TRANSMEM 1026
FT	TRANSMEM 1027
FT	CONFLICT 1037
FT	CONFLICT 303
FT	CONFLICT 372
FT	CONFLICT 372
FT	CONFLICT 385
FT	CONFLICT 461
FT	CONFLICT 461
FT	CONFLICT 665
FT	CONFLICT 665
FT	CONFLICT 763
FT	CONFLICT 763
FT	CONFLICT 775
FT	CONFLICT 775
FT	CONFLICT 778
SQ	SEQUENCE 1037 AA; 113047 MW; 961611E1D24FPAE5 CRC64;
Query Match	47.6%; Score 39; DB 1; Length 1037;
Beat Local Similarity	66.7%; Pred. No. 62;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
Cy 3 RHWLQFAKLTGF 14	
Db 185 RIWIDPAKLNSP 196	
PSRB_VIBAL	STANDARD; PRT; 479 AA.
AC F22825;	
DT 01-AUG-1991 (Rel. 19, Created)	
DT 01-AUG-1991 (Rel. 19, Last sequence update)	
DT 15-JUN-2002 (Rel. 41, Last annotation update)	
DE PHS system, sucrose-specific IIIC component (IIIBC-SCR) (Sucrose-	
DE permease IIIC component) (Phosphotransferase enzyme II, BC component)	
DE (EC 2.7.1.69) (Etl-SCR).	
GN SCRA.	
OS Vibrio alginolyticus.	
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.	
OX NCBI_TaxId=663;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=91071601; PubMed=2174811;	
RA Blatch G.L., Schoille R.R., Woods D.R.;	
RT "Nucleotide sequence and analysis of the Vibrio alginolyticus sucrose"	

FT TRANSMEM 465 485 POTENTIAL.
 FT CONFLICT 404 404 L -> P (IN REF. 2).
 FT CONFLICT 457 457 A -> T (IN REF. 2).
 SQ SEQUENCE 487 AA; 52906 MW; 4D6E8439BEE281D2 CRC64;
 Query Match
 Best Local Similarity 46.3%; Score 38; DB 1; Length 487;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 ID 2 NR1WLOFAKLTG 13
 DB 31 NHTWLYFAVFTG 42

RESULT 7
 TEXT YEAST STANDARD; PRT; 624 AA.
 AC P40090;
 DT 01-FEB-1995 (Rel. 31, Last Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 72.1 kDa protein in FTR1-SPT15 intergenic region.
 GN YER147C.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartwell G., Hunkeler-Smith S.,
 RA Hyman R., Kayser A., Komp C., Laskari D., Lew H., Lin D.,
 RA Meschede D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Borstein D., Davis R.W.,
 RA Submitted (DEC-1994) to the EMBL/Genbank/DBD databases.
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBD databases.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U18917; AAB64674.1; -.
 DR SGD; S0000949; YER147C.
 DR InterPro; IPR000873; AMP-bind.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 KW Hypothetical protein; Ligase.
 SQ SEQUENCE 624 AA; 72140 MW; 1269D931C656E608 CRC64;

Query Match
 Best Local Similarity 46.3%; Score 38; DB 1; Length 624;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 ID 2 NR1WLOFAKLT 11
 DB 461 NKVWLQVETL 470

RESULT 8
 NADM AERPE STANDARD; PRT; 172 AA.
 AC Q9YER8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nicotinamide-nucleotide adenylyltransferase (EC 2.7.7.1) (NAD(+))

DE Pyrophosphorylase (NAD(+)) diphosphorylase (NMN adenylyltransferase).
 GN APO513.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
 CC Desulfurococcaceae; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1.
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Araki A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushiida N., Oguni A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1."
 RL DNA Res. 6:83-101(1999).
 CC -1- CATALYTIC ACTIVITY: ATP + nicotinamide ribonucleotide =
 CC diphosphate + NAD(+).
 CC -1- PATHWAY: NAD BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEAL NMN ADENYLYLTRANSFERASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; APO00059; BAA79478.1; ALT_INIT.
 DR HSSP; Q57961; 1F9A.
 DR InterPro; IPR004821; Cyt. tran rel.
 DR InterPro; IPR004820; Cytidylyltransf.
 DR Pfam; PF01467; Cytidylyltransf. 1.
 DR TIGRPFAM; TIGR00125; Cyt. tran rel. 1.
 KW Transferase; Nucleotidyltransferase; NAD; Complete proteome.
 SQ SEQUENCE 172 AA; 19828 MW; E100C434A2064E4A CRC64;

Query Match
 Best Local Similarity 45.1%; Score 37; DB 1; Length 172;
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 ID 2 NR1WLOFAKLT 11
 DB 82 NKVWVQYLKM 91

RESULT 9
 YGCT_HAEIN STANDARD; PRT; 181 AA.
 ID YGCT_HAEIN
 AC P44097;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein H11036.
 GN H11036.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utechtack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geophagen N.S.M.,
 RA Gnehm C.U., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.,
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RL Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: STRONG. TO E. COLI YGGT.
 CC -1- SIMILARITY: BELONGS TO THE YGGT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U32784; AAC22696.1; ALT_INIT.
 DR TIGR: H11036; -;
 DR InterPro: IPR003425; unk_YGGT.
 DR Pfam: PF02325; YGGT; 2.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 3 23 POTENTIAL.
 FT TRANSMEM 67 87 POTENTIAL.
 FT TRANSMEM 92 112 POTENTIAL.
 FT TRANSMEM 159 179 POTENTIAL.
 SQ SEQUENCE 181 AA; 20200 MW; 3f086a7ea972232 CRC64;
 Query Match 45.1%; Score 37; DB 1; Length 181;
 Best Local Similarity 66.7%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 R1WLOPAKL 11
 DB 24 RAWLOPARY 32

RESULT 10
 PBP5_PSEAE STANDARD; PRT; 310 AA.
 AC P72161;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Penicillin-binding protein 5 precursor (PBP-5) (D-alanyl-D-alanine-
 DE endopeptidase) (EC 3.4.99.-) (DD-endopeptidase).
 GN PBP5 OR PA0869.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RA Song J., Jensen R.A.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 (2)
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RA MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzman L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Bardo L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 CC -1- FUNCTION: CELL WALL FORMATION.

CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S11; ALSO KNOWN AS THE
 CC D-ALANYL-D-ALANINE CARBOXYPEPTIDASE 1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U62582; AAC46341.1; -;
 DR EMBL: AE004521; AAC04258.1; -;
 DR MEROPS: S11.002;
 DR InterPro: IPR001967; Ala/AlaCbpAse1.
 DR Pfam: PF00768; Peptidase_S11; 1.
 DR PRINTS: PR00725; DADCBPTASE1.
 KW Hydrolyase; Peptidoglycan synthesis; Cell wall; Signal; Periplasmic;
 KW Complete proteome.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 310 ACTIATED BY PENICILLIN (BY SIMILARITY).
 FT ACT SITE 66 66 D -> DD (IN REF. 1).
 FT CONFLICT 296 296
 SQ SEQUENCE 310 AA; 34046 MW; B1f62a3a2f9ef641 CRC64;
 Query Match 45.1%; Score 37; DB 1; Length 310;
 Best Local Similarity 56.2%; Pred. No. 38;
 Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 2;

QY 1 NNR1W-LQPAKLGTFT 15
 DB 220 NNR1W1QLTK-TGFT 234

RESULT 11
 OTC_EMENTI STANDARD; PRT; 359 AA.
 ID OTC_EMENTI
 AC P11803;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)
 DE (OTCase) (Ornithine transcarbamylase).
 GN ARGB.
 OS Escherichia nidulans (Aspergillus nidulans).
 OC Bacteria; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; Eurotiella.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87014107; PubMed=3020372;
 RA Upshall A., Gilbert T., Saari G., O'Hara P.J., Weglenski P., Berse B.,
 RA Miller K., Timberlake W.E.;
 RT "Molecular analysis of the argB gene of Aspergillus nidulans.";
 RL Mol. Gen. Genet. 204:349-354(1986).
 CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
 CC + L-citrulline.
 CC -1- PATHWAY: Arginine biosynthesis.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: BELONGS TO THE ATCASP/OTCASP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M29819; AAA50816.1; -;
 DR PIR: S07317; OMANN.

DR HSSP; P00480; 10TH.
 DR InterPro; IPR002029; Asp/Orn_Cotranf.
 DR InterPro; IPR002292; Orn_carbtransf.
 DR Pfam; PF00185; OTCace; 1.
 DR Pfam; PF02729; OTCace.N; 1.
 DR PRINTS; PR00100; AOTCase.
 DR TIGRfams; TIGR00658; orn_carb_tr; 1.
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
 DR TRANSFERASE; Arginine biosynthesis; Mitochondrion; Transit peptide.
 FT TRANSIT 1 24
 FT CHAIN 25 359 ORNITHINE CARBAMOYLTRANSFERASE.
 SQ SEQUENCE 359 AA; 39058 MW; 6B2CBD7BC276F2C4 CRC64;

Query Match 45.1%; Score 37; DB 1; Length 359;
 Best Local Similarity 46.2%; Pred. No. 45;
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 NR1WQFAKTGTF 14
 Db 339 NRLMAAISALEGF 351

RESULT 12

OTC_ASPTB STANDARD; PRT; 361 AA.
 AC 000291;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)
 DE (OTCase) (Ornithine transcarbamylase).
 GN ARG-1 OR ARG.
 OS Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OC NCBI_TaxID=31178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97286514; PubMed=9141661;
 RX Ventura L., Perez-Gonzalez J.A., Ramon D.;
 RT "Cloning and molecular analysis of the Aspergillus terreus arg1 gene
 coding for an ornithine carbamoyltransferase";
 RL FEMS Microbiol. Lett. 149:207-213 (1997).
 CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
 + L-citrulline.
 CC -1- PATHWAY: Arginine biosynthesis.
 CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; Z67741; CAA91554.1; -.
 DR HSSP; P00480; 10TH.
 DR InterPro; IPR002029; Asp/Orn_Cotranf.
 DR InterPro; IPR002292; Orn_carbtransf.
 DR Pfam; PF00185; OTCace; 1.
 DR Pfam; PF02729; OTCace.N; 1.
 DR PRINTS; PR00100; AOTCase.
 DR TIGRfams; TIGR00658; orn_carb_tr; 1.
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
 DR TRANSFERASE; Arginine biosynthesis; Mitochondrion; Transit peptide.
 FT TRANSIT 1 24
 FT CHAIN 25 361 ORNITHINE CARBAMOYLTRANSFERASE.
 SQ SEQUENCE 361 AA; 39129 MW; 9306AD775E499D75 CRC64;

Query Match 45.1%; Score 37; DB 1; Length 361;

Best Local Similarity 46.2%; Pred. No. 45;
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 NR1WQFAKTGTF 14
 Db 341 NRLMAAISALEGF 353

RESULT 13

OTC_ASPTB STANDARD; PRT; 370 AA.
 AC P11066;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)
 DE (OTCase) (Ornithine transcarbamylase).
 GN ARG.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OC NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88167829; PubMed=3443301;
 RX Buxton F.P., Gwynne D.I., Garven S., Sibley S., Davies R.W.;
 RT "Cloning and molecular analysis of the ornithine carbamoyl
 transferase gene of Aspergillus niger";
 RL Gene 61:255-266 (1987).
 CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
 + L-citrulline.
 CC -1- PATHWAY: Arginine biosynthesis.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; M19158; AAA32688.1; -.
 DR PIR; A27362; OMASG.
 DR HSSP; P00480; 10TH.
 DR InterPro; IPR002029; Asp/Orn_Cotranf.
 DR InterPro; IPR002292; Orn_carbtransf.
 DR Pfam; PF00185; OTCace; 1.
 DR Pfam; PF02729; OTCace.N; 1.
 DR PRINTS; PR00100; AOTCase.
 DR TIGRfams; TIGR00658; orn_carb_tr; 1.
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
 DR TRANSFERASE; Arginine biosynthesis; Mitochondrion; Transit peptide.
 FT TRANSIT 1 38
 FT CHAIN 39 370 ORNITHINE CARBAMOYLTRANSFERASE.
 SQ SEQUENCE 370 AA; 39924 MW; 77AAB4B828FA65F CRC64;

Query Match 45.1%; Score 37; DB 1; Length 370;
 Best Local Similarity 46.2%; Pred. No. 46;
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 NR1WQFAKTGTF 14
 Db 350 NRLMAAISALEGF 362

RESULT 14

HMDH_SULSO STANDARD; PRT; 409 AA.
 ID HMDH_SULSO
 AC 008424; OSUWT6;
 DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 3-hydroxy-3-methylglutaryl-coenzyme A reductase (EC 1.1.1.34) (HMG-CoA
 reductase).
 GN HMG-OR S000531 OR C22_020.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 NC NCB1_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=97315236; PubMed=9171410;
 RA Bochar D.A., Brown J.R., Doolittle W.F., Klenk H.-P., Lam W.,
 Schenk M.E., Stauffer C.V., Rodwell V.W.,
 RT "3-hydroxy-3-methylglutaryl coenzyme A reductase of Sulfolobus
 solfataricus: DNA sequence, phylogeny, expression in Escherichia coli
 of the hmgA gene, and purification and kinetic characterization of
 the gene product." J. Bacteriol. 179:3632-3638(1997).
 RL J. Bacteriol. 179:3632-3638(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=20165948; PubMed=10701121;
 RA Charlebois R.L., Singh R.K., Chan-Weiher C.C.-Y., Allard G., Chow C.,
 Confalonieri F., Curtis B., Duguet M., Erasmo G., Faguy D.,
 Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
 Kuehwa N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
 St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
 Ragan M.A., Sengen C.W.,
 RT "Gene content and organization of a 281-kbp contig from the genome of
 the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
 RL Genome 43:116-136(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427226;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 Awayero M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
 De Moers A., Erasmo G., Fletcher C., Gordon P.M.K.,
 Rajkumar-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -1- FUNCTION: CONVERTS HMG-COA TO MEVALONATE.
 CC -1- CATALYTIC ACTIVITY: (R)-mevalonate + CoA + 2 NADP(+) = (S)-3-
 hydroxy-3-methylglutaryl-CoA + 2 NADPH.
 CC -1- PATHWAY: Isoprenoid biosynthesis.
 CC -1- MISCELLANEOUS: STABLE AT 90 DEGREES CELSIUS; OPTIMAL ACTIVITY IS
 AT PH 5.5 AND 85 DEGREES CELSIUS.
 CC -1- SIMILARITY: BELONGS TO THE HMG-COA REDUCTASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in for commercial
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL, U95360; AAC45370.1;
 DR EMBL, Y18930; CAB57768.1;
 DR EMBL, AE006683; AAK40851.1;
 DR InterPro, IPR002202; HMG-CoA_red.
 DR InterPro, IPR004554; HMG-CoA_R_NADP.
 DR Pfam, PF00368; HMG-CoA_red.1.
 DR PRINTS, PRO00071; HMGCOARDPASE.
 DR TIGR, TIGR00533; HMG_COA_R_NADP.1.
 DR PROSITE, PS00066; HMG_COA_REDUCTASE_1; 1.
 DR PROSITE, PS00318; HMG_COA_REDUCTASE_2; 1.
 DR PROSITE, PS01192; HMG_COA_REDUCTASE_3; FALSE_NEG.

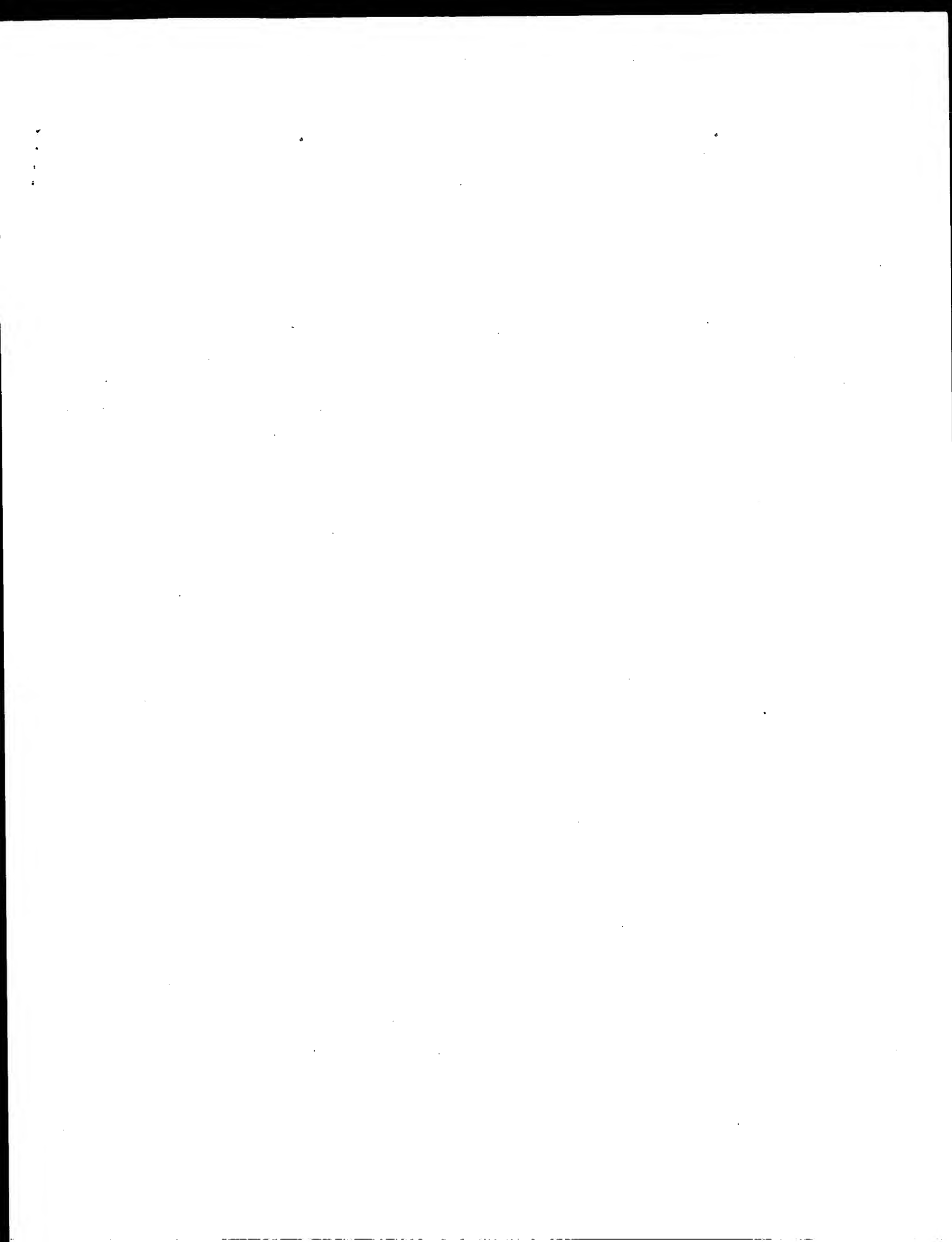
DR PROSITE, PS00065; HMG COA REDUCTASE 4; 1.
 KW Oxidoreductase, Isoprene biosynthesis, NADP, Complete proteome.
 FT ACT SITE 99
 FT ACT SITE 99
 FT ACT SITE 305
 FT ACT SITE 400
 FT ACT SITE 400
 FT CONFLICT 245
 FT CONFLICT 245
 SQ SEQUENCE 409 AA; 44008 MW; 8D1FA64E117CFP46 CRC64;
 Query Match 45.1%; Score 37; DB 1; Length 409;
 Best Local Similarity 50.0%; Pred. No. 52;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 179 NINWLRFPSPRTG 190
 QY 2 NINWLRFPSPRTG 13
 DB 179 NINWLRFPSPRTG 190
 RESULT 15
 ID CALR HUMAN STANDARD; PRT; 490 AA.
 AC P30968; O13941; O14585;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calcitonin receptor precursor (CT-R).
 GN CALR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovarian carcinoma;
 RX MEDLINE=93055406; PubMed=1331173;
 RA Gorn A.H., Lin H.Y., Yamin M., Auron P.E., Flannery M.R.,
 RA Tapp D.R., Manning C.A., Lodish H.F., Krane S.M., Goldring S.R.,
 RT "Cloning, characterization, and expression of a human calcitonin
 receptor from an ovarian carcinoma cell line." J.
 RL J. Clin. Invest. 90:1726-1735(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94359487; PubMed=8078488;
 RA Kuestner R.E., Elrod R.D., Grant F.J., Hagen F.S., Kuiper J.L.,
 RA Mathews S.L., O'Hara P.J., Sheppard P.O., Stroop S.D., Thompson D.L.,
 RA Wilmore T.E., Findlay D.M., Houseman S., Sexton P.M., Moore E.E.,
 RT "Cloning and characterization of an abundant subtype of the human
 calcitonin receptor." J.
 RL Mol. Pharmacol. 46:246-255(1994).
 RN [3]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RC TISSUE=Breast carcinoma;
 RX MEDLINE=96079881; PubMed=7588285;
 RA Albrandt K., Brady E.M.G., Moore C.X., Mull E., Stierzga M.E.,
 RA Beaumont K.,
 RT "Molecular cloning and functional expression of a third isoform of
 the human calcitonin receptor and partial characterization of the
 calcitonin receptor gene." J.
 RL Endocrinology 136:5377-5384(1995).
 RN [4]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RX MEDLINE=94192834; PubMed=8143880;
 RA Frendo J.L., Pichaud F., Delage-Mourroux R., Bouliar Z., Second N.,
 RA Moukhtar M.S., Julienne A.,
 RT "An isoform of the human calcitonin receptor is expressed in TT cells
 and in medullary carcinoma of the thyroid." J.
 RL FEBS Lett. 342:214-216(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Darte M., Gibson A.,
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Breast carcinoma;

```

FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 490 CATIONICIN RECEPTOR.
FT DOMAIN 23 153 EXTRACELLULAR (POTENTIAL) .
FT TRANSMEM 154 173 1 (POTENTIAL) .
FT DOMAIN 174 196 CYTOPLASMIC (POTENTIAL) .
FT TRANSMEM 197 216 2 (POTENTIAL) .
FT DOMAIN 217 236 EXTRACELLULAR (POTENTIAL) .
FT TRANSMEM 237 259 3 (POTENTIAL) .
FT DOMAIN 260 276 CYTOPLASMIC (POTENTIAL) .
FT TRANSMEM 277 296 4 (POTENTIAL) .
FT DOMAIN 297 312 EXTRACELLULAR (POTENTIAL) .
FT TRANSMEM 313 332 5 (POTENTIAL) .
FT DOMAIN 337 359 CYTOPLASMIC (POTENTIAL) .
FT TRANSMEM 360 377 6 (POTENTIAL) .
FT DOMAIN 378 389 EXTRACELLULAR (POTENTIAL) .
FT TRANSMEM 390 411 7 (POTENTIAL) .
FT DOMAIN 412 490 CYTOPLASMIC (POTENTIAL) .
FT DOMAIN 444 449 POLY-ALA.
FT CARBOHYD 28 28 N-LINKED.
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARSPLIC 1 47 MISSING (IN ISOFORM 3) .
FT VARSPLIC 175 190 MISSING (IN ISOFORM 2 AND ISOFORM 3) .
FT VARSPLIC 463 463 P -> L.
FT CONFLICT 363 363 /FTID=VAR. 003580.
FT SEQUENCE 490 AA; 363 363 I -> T (IN REF. 3; AAC50301) .
SO 490 AA; 363 363 38ECTAA4A9F07FF5D CRC64;

Query Match 45.1%; Score 37; DB 1; Length 490;
Best Local Similarity 40.0%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 NNRIMLOFAKLGT 15
Db 124 NNRITMSNYTMCAFT 138

```



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 12.6711 Seconds

(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524d-100

Perfect score: 82

Sequence: 1 NNRIWLPFAKLTGFT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_yeast:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	63.4	507	10	Q9FY19
2	47	57.3	457	10	Q9ZUB7
3	47	57.3	1161	10	Q9LOD1
4	44	53.7	281	17	Q97OD6
5	43	52.4	466	10	Q9CAL5
6	42	51.2	266	16	Q8YQJ9
7	42	51.2	385	3	Q9C109
8	42	51.2	439	2	Q93061
9	42	51.2	474	10	Q9W4S2
10	42	51.2	475	10	Q94EH4
11	42	51.2	475	10	Q949Z1
12	42	51.2	492	10	Q9LNG3
13	42	51.2	538	12	Q9FX25
14	42	51.2	540	10	Q92699
15	42	51.2	853	3	Q12440
16	41	50.0	220	5	Q9U2G1

17	41	50.0	235	16	Q8XET9	Q8XET9 escherichia
18	41	50.0	452	16	Q34521	Q34521 bacillus su
19	40	48.8	75	2	Q9FCW9	Q9FCW9 escherichia
20	40	48.8	246	17	Q9Y9C6	Q9Y9C6 aeropyrum p
21	40	48.8	297	5	Q9W106	Q9W106 drosophila
22	40	48.8	309	5	Q9VAF2	Q9VAF2 drosophila
23	40	48.8	367	10	Q9SR79	Q9SR79 arabidopsis
24	40	48.8	382	10	Q8SP99	Q8SP99 oryza sativ
25	40	48.8	419	10	Q9LS84	Q9LS84 arabidopsis
26	40	48.8	419	10	Q93Z21	Q93Z21 arabidopsis
27	40	48.8	756	16	P72877	P72877 synechocyst
28	40	48.8	1049	16	Q8Y3H0	Q8Y3H0 ralsonia s
29	40	48.8	1501	3	Q9C2Y4	Q9C2Y4 magnaporthe
30	40	48.8	1501	3	Q9C2Y4	Q9C2Y4 magnaporthe
31	39	47.6	138	10	Q9FTH4	Q9FTH4 arabidopsis
32	39	47.6	217	5	Q95UV7	Q95UV7 blomphalari
33	39	47.6	217	16	Q8YFT0	Q8YFT0 bruceella me
34	39	47.6	225	3	Q14262	Q14262 schizosacch
35	39	47.6	376	2	Q9RG16	Q9RG16 bacteroides
36	39	47.6	399	2	Q9FA51	Q9FA51 rhododactyl
37	39	47.6	485	16	Q9KVV3	Q9KVV3 vibrio chol
38	39	47.6	677	5	Q9XTV1	Q9XTV1 caenorhabdi
39	39	47.6	711	10	Q9SCB2	Q9SCB2 eremoseyne p
40	39	47.6	1037	16	Q8ZNT7	Q8ZNT7 salmonella
41	39	47.6	1037	16	Q8ZAS4	Q8ZAS4 salmonella
42	39	47.6	1037	16	Q8XEH2	Q8XEH2 escherichia
43	39	47.6	1277	12	Q9UGJ7	Q9UGJ7 sorghum chl
44	38.5	47.0	652	5	Q8SSV3	Q8SSV3 dictyosteli
45	38	46.3	145	8	Q20657	Q20657 haplochromi

ALIGNMENTS

RESULT 1	Q9FY19	PRELIMINARY;	PRT;	507 AA.
ID	Q9FY19			
AC	Q9FY19			
DC	01-MAR-2001 (TREMREL. 16, Created)			
DT	01-MAR-2001 (TREMREL. 16, Last sequence update)			
DT	01-JUN-2002 (TREMREL. 21, Last annotation update)			
DE	Pollen major allergen 2 protein precursor.			
GN	JNA2.			
OS	Juniperus ashei (Ozark white cedar).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.			
OX	NCBI_TaxID=13101;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=MALE POLLEN;			
RX	MEDLINE=20403896; PubMed=10944464;			
RA	Yokoyama M., Miyahara M., Shimizu K., Kino K., Tsunoo H.;			
RT	"Purification, identification and cDNA cloning of Jun a 2, the second			
RT	major allergen of mountain cedar pollen."			
RL	Biochem. Biophys. Res. Commun. 275:195-202(2000).			
CC	-1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES			
CC	(POLYGALACTURONASES).			
CC	EMBL: AJ04653; CAC05582.1; -			
DR	HSSP; P26509; 1BHE.			
DR	InterPro; IPR000743; GH28.			
DR	InterPro; IPR000408; Reg_cht condens.			
DR	Pfam; PF00295; Glyco_hydro_28; 1.			
DR	PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.			
DR	PROSITE; PS00626; RCCL_2; UNKNOWN_1.			
KW	Cell wall; Glycosidase; Hydrolase; Signal.			
FT	SIGNAL 1			
FT	POTENTIAL.			
SQ	SEQUENCE 507 AA; 55730 MW; 2B2E0A5E58FE5A CRC64;			
Query Match	63.4%;	Score 52;	DB 10;	Length 507;
Best Local Similarity	57.1%;	Pred. No. 0.75;		
Matches	8;	Conservative	4;	Mismatches 2;
		Indels	0;	Gaps 0;
Qy	1	NNRIWLPFAKLTGFT 14		

RA Kawarayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kuehida N., Oguchi A.,
 RA Aoki K.-I., Maeda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.,
 RT "Complete genome sequence of an aerobic thermophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain 7.",
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AF000987; BAB6737.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 281 AA; 30125 MW; 465BR45797F383DC CRC64;

Query Match 53.7%; Score 44; DB 17; Length 281;
 Best Local Similarity 57.1%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 DB 243 NNTVITQFPPLTGF 256

RESULT 5
 ID 09CAL5 PRELIMINARY; PRT; 468 AA.
 AC 09CAL5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative polygalacturonase.
 GN F24013.7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin B.,
 RA Kim C.-U., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luyos J.S., Malt R., Marziani A.,
 RA Miltner J., Miranda M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana."
 RL Nature 408:816-820(2000).
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYGALACTURONASES).
 DR EMBL; AC010796; AAG52465.1; -
 DR InterPro; IPR000886; ER_target.
 DR InterPro; IPR000743; GH28.
 DR Pfam; PF00285; Glyco_hydro_28.1.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 KW Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 468 AA; 50900 MW; B92D6E5168D7455 CRC64;

Query Match 52.4%; Score 43; DB 10; Length 468;
 Best Local Similarity 72.7%; Pred. No. 27;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RIMLOFAKLTG 13
 DB 147 RNMLOFSLQ 157

RESULT 6
 ID 08Y0J9 PRELIMINARY; PRT; 266 AA.
 AC 08Y0J9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein A13824.
 GN A13824.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OC NCBI_TaxId=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120."
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AP003594; BAB75523.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 266 AA; 30887 MW; 741814848A5B8C0 CRC64;

Query Match 51.2%; Score 42; DB 16; Length 266;
 Best Local Similarity 58.3%; Pred. No. 22;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 RIMLOFAKLTG 13
 DB 206 RNMLOFSLQ 217

RESULT 7
 ID 09C1O9 PRELIMINARY; PRT; 385 AA.
 AC 09C1O9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Pheromone receptor CPRalp.
 GN CPRAL.
 OS Cryptococcus neoformans (Filobasidiella neoformans).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OC NCBI_TaxId=5207;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chang Y.C., Kwon-Chung K.J.;
 RT "Pheromone receptor of Cryptococcus neoformans mating type a."
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF250141; AAK31936.1; -
 DR InterPro; IPR001499; STE3_GPCR.
 DR Pfam; PF02076; STE3_1.
 DR PRINTS; PR00899; GPCRSTE3.
 KW Receptor.
 SQ SEQUENCE 385 AA; 43122 MW; 146E23DBC9259E8 CRC64;

Query Match 51.2%; Score 42; DB 3; Length 385;
 Best Local Similarity 63.6%; Pred. No. 33;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 4 IWLQFAKLTG 14

Db 37 VWLTFANLQGF 47

RESULT 8

Q93061 PRELIMINARY; PRT; 439 AA.
 AC Q93061;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Integral membrane protein.
 GN YIAA2.
 OS Klebsiella oxytoca.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Klebsiella.
 OC NCBI_TaxID=571,
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M5A1;
 RA Darois V.A., Menge K.L., Vande Weghe J.;
 RT "The ascorbic acid degradation pathway of Klebsiella oxytoca M5a1."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF282849; AKK69524.1;
 DR InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00083; sugar tr; 1.
 SQ SEQUENCE 439 AA; 47463 MW; AB37CDBB64967257 CRC64;

Query Match 51.2%; Score 42; DB 2; Length 439;
 Best Local Similarity 58.3%; Pred. No. 37;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 NR1WLOFAKLTG 13
 DB 330 NQ1WLSYALVG 341

RESULT 9

Q8W4S2 PRELIMINARY; PRT; 474 AA.
 AC Q8W4S2;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE At1g60590/F8A5.12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bower L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY060568; AAL31197.1;
 DR InterPro; IPR000743; GH28.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.
 SQ SEQUENCE 474 AA; 51964 MW; 9631844B83E209C6 CRC64;

Query Match 51.2%; Score 42; DB 10; Length 474;
 Best Local Similarity 54.5%; Pred. No. 40;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 WLOFAKLTGFT 15

Db 139 WIEFTKLGIT 149

RESULT 10

Q94EH4 PRELIMINARY; PRT; 475 AA.
 AC Q94EH4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE At1g48100/F21D18.17.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bower L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYGALACTURONASES).
 DR EMBL; AF410319; AAK5305.1;
 DR InterPro; IPR000743; GH28.
 DR InterPro; IPR000408; Reg_chr_condens.
 DR InterPro; IPR001412; rrm_syt1.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
 DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.
 DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
 KW Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 475 AA; 51380 MW; 4A42C4355827D30A CRC64;

Query Match 51.2%; Score 42; DB 10; Length 475;
 Best Local Similarity 53.3%; Pred. No. 41;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NR1WLOFAKLTGFT 15

DB 159 NKNQWLTVFYRLDGT 173

RESULT 11

Q949Z1 PRELIMINARY; PRT; 475 AA.
 AC Q949Z1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative polygalacturonase pgl.
 GN F21D18.18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
 RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bower L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B.,
 RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Tracy S.E.,

RA Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F21D18.18 (GI:8778527).";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 DR EMBL; AY050798; AK92733.1; -
 DR InterPro; IPR000743; GH28.
 DR InterPro; IPR00408; Reg_ch_r-condens.
 DR InterPro; IPR001412; CRNA-synt.1.
 DR Pfam; PF00295; Glyco_hydro.28; 1.
 DR PROSITE; PS00178; AA-TRNA LIGASE I; UNKNOWN_1.
 DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.
 DR PROSITE; PS00526; RCC1_2; UNKNOWN_1.
 KM Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 475 AA; 51352 MW; ACB066CF725F2C8F CRC64;
 Query Match 51.2%; Score 42; DB 10; Length 475;
 Best Local Similarity 53.3%; Pred. No. 41;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 NNRIMLOPAKLTGFT 15
 DB 159 NKNQWLVFRLDGT 173
 RESULT 12
 ID 09LNG3 PRELIMINARY; PRT; 492 AA.
 AC 09LNG3;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE F21D18.18;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bei Q., Chin C., Chiu J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lanz C., Li J., Liu A., Liu K., Liu S., Mukharasy N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F21D18 from chromosome
 I.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiu J., Choi E.,
 RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lanz C., Li J., Liu A., Liu K., Liu S., Mukharasy N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.R.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 DR EMBL; AC023673; AF79535.1; -
 DR InterPro; IPR000743; GH28.
 DR InterPro; IPR00408; Reg_ch_r-condens.

DR InterPro; IPR001412; CRNA-synt.1.
 DR Pfam; PF00295; Glyco_hydro.28; 1.
 DR PROSITE; PS00178; AA-TRNA LIGASE I; UNKNOWN_1.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 DR PROSITE; PS00526; RCC1_2; UNKNOWN_1.
 KM Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 492 AA; 53395 MW; 42BAF98B9E30AC1A CRC64;
 Query Match 51.2%; Score 42; DB 10; Length 492;
 Best Local Similarity 53.3%; Pred. No. 42;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 NNRIMLOPAKLTGFT 15
 DB 159 NKNQWLVFRLDGT 173
 RESULT 13
 ID 09PX25 PRELIMINARY; PRT; 538 AA.
 AC 09PX25;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE Nonstructural protein.
 GN ORF-ALPHA.
 OS Periplaneta fuliginosa densovirus.
 OC Viruses; ssDNA viruses; Parvoviridae; Densovirinae; Densovirus.
 NCBI_TaxID=97344;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Yamagishi J., Hu Y., Zhang J., Bando H.;
 RT "Genome organization and mRNA structure of Periplaneta fuliginosa
 RT densovirus imply alternative splicing involvement in viral gene
 RT expression.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Guo H.T., Zhang J.M., Hu Y.Y.;
 RT "Complete sequence of periplaneta fuliginosa parvovirus.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB028936; BA082961.1; -
 DR EMBL; AF192260; AAF04297.1; -
 KW Nonstructural protein.
 SQ SEQUENCE 538 AA; 61746 MW; 6E168FA9236DC5 CRC64;
 Query Match 51.2%; Score 42; DB 12; Length 538;
 Best Local Similarity 61.5%; Pred. No. 46;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NNRIMLOPAKLTG 13
 DB 181 SNQWLOPAKVG 193
 RESULT 14
 ID 022699 PRELIMINARY; PRT; 540 AA.
 AC 022699;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative polygalacturonase.
 GN F8A5.12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Federspiel N.A., Palm C.J., Conway A.B., Kuttz D.B., Conway A.R.,
 RA Federspiel N.A., Buehler E., Dewar K., Feng J., Kim C.,
 RA Au M., Araujo R., Brendel V., Buehler E.,

RA Li Y., Oji O., Osborne B.I., Shin P., Sun H., Toriumi M.,
 RA Vysotskaya V.S., Yu G., Ecker J., Theologis A., Davis R.W.;
 RT "Genomic sequence of Arabidopsis";
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 CC EMBL: AC002292; AAB71972.1; -.
 DR InterPro: IPR000743; GH28.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 KW Cell wall; Glycosidase; Hydrolase
 SQ SEQUENCE 540 AA; 59631 MW; 532B79A84EBB904 CRC64;
 QY 5 WLOPAKLTGFT 15
 DB 205 WIEFTKLSGIT 215
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Query Match 51.2%; Score 42; DB 10; Length 540;
 Best Local Similarity 54.5%; Pred. No. 46;
 RESULT 15
 ID Q12440 PRELIMINARY; PRT; 853 AA.
 AC Q12440;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE ORF YLR127C.
 GN APC2 OR L3108 OR L3105 OR L9233.13 OR YLR127C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxID=4932;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FY23/RD005;
 RA Verhaesselt P.; Volckaert G.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Delius H.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Delius H.; Hebling U.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MIPs;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Pauley A.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Johnston M.; Andrews S.; Brinkman R.; Cooper J.; Ding H.; Du Z.;
 RA Favello A.; Fulton L.; Gattung S.; Greco T.; Kirsten J.; Kucaba T.;
 RA Hallsworth K.; Hawkins J.; Hillier L.; Jier M.; Johnson D.;
 RA Johnston L.; Langston Y.; Latreille P.; Le T.; Mardis E.; Menezes S.;
 RA Miller N.; Nhan M.; Pauley A.; Peluso D.; Rifken L.; Riles L.;
 RA Taich A.; Trevasakis E.; Vignati D.; Wilcox L.; Wohlman P.; Vaudin M.;
 RA Wilson R.; Waterston R.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Waterston R.;

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X89514; CAA61705.1; -.
 DR EMBL: X91258; CAA62638.1; -.
 DR EMBL: Z73299; CAA97696.1; -.
 DR EMBL: U53877; AAB82373.1; -.
 DR EMBL: Z73300; CAA97698.1; -.
 DR SGD: S0004117; APC2.
 DR InterPro: IPR001373; Cullin.
 DR SMART: SM00182; CULLIN; 1.
 DR PROSITE: PS50069; CULLIN_2; 1.
 SQ SEQUENCE 853 AA; 99977 MW; E559A0E28C784DE2 CRC64;
 QY 2 NR1WLOPAKLTG 13
 DB 248 NQYWSQFSKLVG 259
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Query Match 51.2%; Score 42; DB 3; Length 853;
 Best Local Similarity 58.3%; Pred. No. 74;
 Search completed: April 20, 2003, 13:13:12
 Job time: 14.6711 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds

(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524d-101

Perfect score: 76

Sequence: 1 LQFAKLTGFTLMGKG 15

Scoring table: BLOSUM62
Gapop: 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	514	S48730	Cry j II protein -
2	76	100.0	514	JC2498	second major aller
3	54	71.1	514	JC7100	polygalacturonase
4	48	63.2	507	JC7366	un a 2 protein -
5	45	59.2	540	B96631	probable polygalac
6	44	57.9	492	C96821	protein F21D18.18
7	43	56.6	194	T35332	hypothetical prote
8	43	56.6	502	H75290	hypothetical prote
9	43	56.6	503	T29870	hypothetical prote
10	43	56.6	519	A96542	probable cytochrom
11	43	56.6	161	B86368	protein F28C11.9
12	43	53.9	468	H96728	probable polygalac
13	41	52.6	217	AB3416	hypothetical membr
14	40	52.6	231	S50517	hypothetical prote
15	40	52.6	285	I65760	alcohol sulfotrans
16	40	52.6	353	F72394	glutamate 5-kinase
17	40	52.6	358	C72626	hypothetical prote
18	40	52.6	371	F69429	asparagine synthet
19	40	52.6	435	T48638	polygalacturonase
20	40	52.6	567	S67702	ABC transporter si
21	39	51.3	217	AB3432	thiol:disulfide in
22	39	51.3	359	B40354	arylalkylphospha
23	39	51.3	382	D84856	hypothetical prote
24	39	51.3	542	F81075	protein F20B24.8
25	39	51.3	572	F81075	arginyl-tRNA synth
26	39	51.3	572	F81075	arginyl-tRNA synth
27	39	51.3	591	B83457	glyoxylate carboxi
28	39	51.3	1020	B82604	cation efflux syst
29	38	50.0	230	D70718	hypothetical prote

30	38	50.0	312	2	T05348	hypothetical prote
31	38	50.0	329	2	B90879	fermentative D-lac
32	38	50.0	329	2	H85739	fermentative D-lac
33	38	50.0	329	2	G64888	D-lactate dehydrog
34	38	50.0	369	2	S68692	deoxyhypusine synt
35	38	50.0	443	2	T46127	hypothetical prote
36	38	50.0	461	2	S74414	hypothetical prote
37	38	50.0	469	2	G90422	Na+/H+-exchanging
38	38	50.0	534	2	T34855	sugar transport pr
39	38	50.0	639	2	AB3456	probable glyoxylat
40	38	50.0	2793	2	B90784	long-chain-fatty-a
41	38	50.0	2806	2	D85644	hypothetical prote
42	37.5	49.3	244	2	F69830	hypothetical prote
43	37.5	49.3	466	2	T02324	conserved hypotne
44	37.5	49.3	1679	2	S49802	hypothetical prote
45	37	48.7	224	2	T11207	probable membrane
						spap protein - Sal

ALIGNMENTS

RESULT 1

S48730
Cry j II protein - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)

C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999

C:Accession: S48730

R:Namba, M.; Kurose, M.; Torijoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.;

FEBS Lett. 353, 124-128, 1994

A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese ceda

A:Reference number: S48730; MUID:95010777; PMID:7926035

A:Accession: S48730

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-514 <NAM>

A:Cross-references: GB:D37765; NID:9577695; PIDN:BA07021.1; PID:di007598; PID:957769

Query Match 100.0%; Score 76; DB 2; Length 514;

Best Local Similarity 100.0%; Pred No. 7.8e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LQFAKLTGFTLMGKG 15

Db 140 LQFAKLTGFTLMGKG 154

RESULT 2

JC2498
second major allergen Cry j II precursor - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)

C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000

C:Accession: JC2498; PC2346; A60147

R:Komiya, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994

A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese

A:Reference number: JC2498; MUID:94271186; PMID:8002972

A:Accession: JC2498

A:Molecule type: mRNA

A:Residues: 1-514 <KOM>

A:Cross-references: DDBJ:D29772; NID:9506857; PIDN:BA06172.1; PID:9506858

A:Accession: PC2346

A:Molecule type: protein

A:Residues: 52-61 <KO2>

R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matuhasi, T.

Allergy 45, 309-312, 1990

A:Title: Identification of the second major allergen of Japanese cedar pollen.

A:Reference number: A60147; MUID:90342988; PMID:2382797

A:Accession: A60147

A:Molecule type: protein

A:Residues: 55-64 <SAK>

C:Keywords: glycoprotein; pollen

F:1-54/Domains: signal sequence

F:55-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 7,8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGKG 15
|||
Db 140 LQFAKLTGFTLMGKG 154

RESULT 3

UC7100 Polygalacturonase Cha o 2 - Japanese cypress

C/Species: Chamaecyparis obtusa (Japanese cypress)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000

C/Accession: JC7100; PC7026

R/Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.

Biochem. Biophys. Res. Commun. 263, 166-171, 1999

A/Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all

A/Reference number: JC7100; MUID:99417540; PMID:10486272

A/Accession: JC7100

A/Molecule type: mRNA

A/Residues: 1-514 <MOR>

A/Accession: PC7026

A/Molecule type: protein

A/Residues: 51-62 <MO2>

Query Match 71.1%; Score 54; DB 2; Length 514;
Best Local Similarity 73.3%; Pred. No. 0.081;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGKG 15
|||
Db 140 LQFAKLTGFTLMGKG 154

RESULT 4

UC7366 Jun a 2 protein - mountain cedar

C/Species: Juniperus ashei (mountain cedar)
C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000

C/Accession: JC7366; PC7093

R/Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.

Biochem. Biophys. Res. Commun. 275, 195-202, 2000

A/Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all

A/Reference number: JC7366

A/Accession: JC7366

A/Molecule type: mRNA

A/Residues: 1-507 <YOK>

A/Cross-references: GB:AJ04653

A/Accession: PC7093

A/Molecule type: protein

A/Residues: 55-63 <Y02>

C/Comment: This protein, a second major allergen of mountain cedar pollen, which is inv

C/Keywords: glycoprotein; pollen

Query Match 63.2%; Score 48; DB 2; Length 507;
Best Local Similarity 60.0%; Pred. No. 0.99;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGKG 15
|||
Db 141 MHPARLTDPMIMGKG 155

RESULT 5

B9631 Probable polygalacturonase P8A5.12 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: B9631

R/Theologis, A.; Ecker, J.R.; Palm, C.D.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shum, P.; Southwick, A.M.; Sun, H.; Tallon

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: B9631

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-540 <STO>

A/Cross-references: GB:AE005173; NID:92462753; PIDN:AAW71972.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

QY 1 LQFAKLTGFTLMGKG 15
|||
Db 206 IEFKLSGITIGGNG 220

RESULT 6

C96521 Hypochemical protein SC5H.16 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C/Accession: C96521

R/Theologis, A.; Ecker, J.R.; Palm, C.D.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shum, P.; Southwick, A.M.; Sun, H.; Tallon

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: C96521

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-492 <STO>

A/Cross-references: GB:AE005173; NID:98778527; PIDN:AAF79535.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

Query Match 57.9%; Score 44; DB 2; Length 492;
Best Local Similarity 60.0%; Pred. No. 5.2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGKG 15
|||
Db 164 LVFRLDGFPSGKG 178

RESULT 7

T35332 Hypochemical protein SC5H.16 - Streptomyces coelicolor

C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

C/Accession: T35332

R/Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1999

A/Reference number: Z21575

A:Accession: T35332
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-194 <OLI>
A:Cross-references: EMBL:AL049863; PIDD:CA642941.1; GSPDB:GN00070, SCOEDB:SCSH1.16
A:Experimental source: strain A3(2)
A:Genetics:
A:Gene: SCOEDB:SCSH1.16

Query Match 56.6%; Score 43; DB 2; Length 194;
Best Local Similarity 63.6%; Pred. No. 2.9;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 KUTGFTLMGK 15
DB 37 RITGFTLMGK 47

RESULT 8
H75290
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: H75290

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickley, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Vamathevan, J.U.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75250

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-502 <WHI>
A:Cross-references: GB:AE002061; GB:AE000513; NID:96460095; PIDD:AAF11844.1; PID:9646010
A:Experimental source: strain R1
A:Genetics:
A:Gene: DR2286
A:Map position: 1

Query Match 56.6%; Score 43; DB 2; Length 502;
Best Local Similarity 47.1%; Pred. No. 8;
Matches 8; Conservative 7; Mismatches 0; Indels 2; Gaps 1;

QY 1 LQPAKLTGF--TLMGK 15
DB 287 VEPARLGGYRLTLGK 303

RESULT 9
T29870
hypothetical protein F32B5.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T29870

R:Bedwith, J.; Graves, T.; Biewald, T.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid F32B5.
A:Reference number: Z20702
A:Accession: T29870
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1503 <LED>
A:Cross-references: EMBL:AF003148; PIDD:AA54212.1; GSPDB:GN00019; CESP:F32B5.6
A:Experimental source: strain Bristol N2; clone F32B5
A:Genetics:
A:Gene: CESP:F32B5.6
A:Map position: 1
A:Insertions: 34/2; 71/1; 210/3; 249/3; 299/3; 358/2; 398/2; 415/1; 453/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F32B5.6

Query Match 56.6%; Score 43; DB 2; Length 503;
Best Local Similarity 53.8%; Pred. No. 8.1;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 2 QPAKLTGFTLMGK 14
DB 429 KFAEFAAGFTVLGK 441

RESULT 10

A96542
probable cytochrome P450 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 20-Apr-2001
C:Accession: A96542

R:Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A96542

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-519 <STO>
A:Cross-references: GB:AE005173; NID:911054637; PIDD:AA627882.1; GSPDB:GN00141
A:Genetics:
A:Gene: F1706.8
A:Map position: 1

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:58/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 56.6%; Score 43; DB 2; Length 519;
Best Local Similarity 61.5%; Pred. No. 8.3;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 QPAKLTGFTLMGK 14
DB 239 QFSKLWGISLFGK 251

RESULT 11

B86368
protein F28C11.9 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B86368

R:Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86368

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1161 <STO>
A:Cross-references: GB:AE005172; NID:98778576; PIDD:AAF79584.1; GSPDB:GN00141
A:Genetics:
A:Gene: F28C11.9
A:Map position: 1

Query Match 56.6%; Score 43; DB 2; Length 1161;
Best Local Similarity 53.3%; Pred. No. 20;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 LQFALGTGTLMGKG 15
 ||:|||||
 Db 827 LQFSKLGVPQGGK 841

RESULT 12

H96728

probable polygalacturonase F24J13.7 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: H96728

R:Theologas, A.; Becker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzilli,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H96728

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-468 <STO>

A:Cross-references: GB:AE005173; NID:g6175142; PIDN:AAF04869.1; GSPDB:GN00141

C:Genetics:

A:Gene: F24J13.7

A:Map position: 1

Query Match 53.9%; Score 41; DB 2; Length 468;

Best Local Similarity 53.3%; Pred. No. 17;

Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 LQFALGTGTLMGKG 15

||:|||||

Db 150 LQFSKLGVPQGGK 164

RESULT 13

AB3416

hypothetical membrane spanning protein BMEI1312 [imported] - Brucella melitensis (strain

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AB3416

R:DeVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AB3416

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-217 <KUR>

A:Cross-references: GB:AE008917; PIDN:AA152493.1; PID:gl7983302; GSPDB:GN00130

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI1312

A:Map position: 1

Query Match 52.6%; Score 40; DB 2; Length 217;

Best Local Similarity 69.2%; Pred. No. 12;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 FAKLTGFTLMGKG 15

|||||

Db 125 FALMGFTLMGKG 137

RESULT 14

SS0517

hypothetical protein YEL072w - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 19-Apr-2002

C:Accession: SS0517

R:Dietrich, F.S.

submitted to the EMBL Data Library, December 1994

A:Description: The sequence of S. cerevisiae cosmid 9669, 8334, 8199, and lambda clone

A:Reference number: SS0434

A:Accession: SS0517

A:Molecule type: DNA

A:Residues: 1-231 <DIE>

A:Cross-references: EMBL:U18795; NID:g603241; PID:g603246; GSPDB:GN00005; MIPS:YEL072w

C:Genetics:

A:Gene: MIPS:YEL072w

A:Cross-references: GSD:S0000798

A:Map position: 5L

Query Match 52.6%; Score 40; DB 2; Length 231;

Best Local Similarity 63.6%; Pred. No. 12;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 KLTGFTLMGKG 15

|||||

Db 137 EITGFTSMGKG 147

RESULT 15

165760

alcohol sulfotransferase (EC 2.8.2.2) - rat

N:Alternate names: hydroxysteroid sulfotransferase; senescence marker protein 2B

C:Species: Rattus norvegicus (Norway rat)

C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 18-Jun-1999

C:Accession: 165760; 152407

R:Macade, T.; Ogura, K.; Satsukawa, M.; Okuda, H.; Hiratsuka, A.

Chem. Biol. Interact. 92, 87-105, 1994

A:Title: Molecular cloning and functions of rat liver hydroxysteroid sulfotransferases

A:Reference number: 152849; MUID:94306585; PMID:8033273

A:Accession: 165760

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-285 <RBS>

A:Cross-references: GB:D14989; NID:g303801; PIDN:BA03634.1; PID:g303802

R:Song, C.S.; Kim, J.M.; Roy, A.K.; Chatterjee, B.

Biochemistry 29, 542-551, 1990

A:Title: Structure and regulation of the senescence marker 2 gene promoter gene promo:

A:Reference number: 152408; MUID:90148982; PMID:2302387

A:Accession: 152407

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 2-68 <RBS>

A:Cross-references: GB:M29302; NID:g206995; PIDN:AAA42152.1; PID:g206996

C:Genetics:

A:Inserts: 46/1

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Query Match 52.6%; Score 40; DB 2; Length 285;

Best Local Similarity 80.0%; Pred. No. 16;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 LTGFTLMGKG 15

|||||

Db 240 LTGFTPMRKG 249

Search completed: April 20, 2003, 13:15:52

Job time : 7.07895 secs

GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524d-101

Sequence: 1 LQPAKLTGFTLMGKG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	514	MPA2_CRYVA	P43212 cryptomeria
2	40	52.6	231	YE12_YEAST	P39975 saccharomyc
3	40	52.6	285	YEH3_RAT	P50235 ratulus notv
4	40	52.6	353	PROB_THEMA	Q9wyd0 thermotoga
5	40	52.6	567	Y095_SYNY3	Q55884 synectocyst
6	39	51.3	358	PONI_RABIT	P27170 mycotoxagus
7	38	50.0	230	Y966_MYCTU	P71544 mycobacteri
8	38	50.0	329	LDHD_ECOLI	P56643 escherichia
9	38	50.0	369	DHVS_HUMAN	P48366 homo sapien
10	38	50.0	759	RRP3_TAKTE	P18884 influenza a
11	37.5	49.3	1679	YMF9_YEAST	Q04958 saccharomyc
12	37	48.7	224	SPAP_SALIT	P40700 salmonella
13	37	48.7	224	SPAP_SALIT	P40700 salmonella
14	37	48.7	244	CYBH_ALCEU	P31898 alcalligenes
15	37	48.7	280	LEPA_CHMU	O9f111 chlamydia m
16	37	48.7	280	LEPA_CHMU	O84536 chlamydia t
17	37	48.7	287	Y193_HAETN	Q57427 haemophilus
18	37	48.7	397	Y664_MYCTU	Q11055 mycobacteri
19	37	48.7	407	MOSC_RHIME	Q07609 rhizobium m
20	37	48.7	438	DOCTA_BRAJA	O9anr2 bradyrhizob
21	37	48.7	911	TB11_NEIMB	O01995 neisseria m
22	37	48.7	915	TB11_NEIMB	O01995 neisseria s
23	36	47.4	160	NRDI_STRPY	Q98xx2 streptococ
24	36	47.4	210	NRDI_STRPY	Q98xx2 streptococ
25	36	47.4	230	DOCTA_BRAJA	O05406 bacillus su
26	36	47.4	315	YB52_SYNY3	Q01958 entamoeba h
27	36	47.4	331	YB52_SYNY3	P74221 synectocyst
28	36	47.4	347	ADH2_KLAMA	O94942 kluyveromyc
29	36	47.4	608	YLPK_CAREL	P33385 caenorhabdi
30	36	47.4	611	GLMS_NETMA	O91wn3 n glucosami
31	36	47.4	611	GLMS_NETMA	O91wn3 n glucosami
32	36	47.4	611	Y043_MENJA	O60348 methanococ
33	36	47.4	810	YP07_YEAST	O08960 saccharomyc

ALIGNMENTS

RESULT ID	MPA2_CRYVA	STANDARD	PRT	514 AA.
AC	P43212;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)			
DE	Major pollen allergen Cry j 2 (Cry j II).			
OS	Cryptomeria japonica (Japanese cedar)			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.			
OX	NCBI_Taxid=3369;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=Pollen;			
RX	MEDLINE=9501077; PubMed=7926035;			
RA	Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,			
RA	Usui M., Kurimoto M.,			
RT	"Molecular cloning of the second major allergen, Cry j II, from			
RT	Japanese cedar pollen."			
RL	FEBS Lett. 353:124-128 (1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pollen;			
RX	MEDLINE=94271186; PubMed=8002972;			
RA	Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.,			
RT	"CDNA cloning and expression of Cry j II the second major allergen of			
RT	Japanese cedar pollen."			
RL	Biochem. Biophys. Res. Commun. 201:1021-1028 (1994).			
RN	[3]			
RP	SEQUENCE OF 55-64.			
RC	MEDLINE=90342988; PubMed=2382797;			
RA	Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matunasi T.,			
RT	"Identification of the second major allergen of Japanese cedar			
RT	pollen."			
RL	Allergy 45:309-312 (1990).			
CC	-1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-			
CC	galactosiduronic linkages in pectate and other galacturonans.			
CC	-1- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).			
CC	-1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES			
CC	(POLYGALACTURONASES).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; D37765; BAA07021.1; -			
DR	EMBL; D29772; BAA06172.1; -			
DR	HSSP; P26509; 1BHE.			
DR	InterPro; IPR000743; GH28.			
DR	Pfam; PF00295; Glyco_hydro_28; 1.			

DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 KM Hydrolyase; Glycosidase; Cell wall; signal; Zymogen; Fruit ripening;
 KM Amyloplast; Glycoprotein; Allergen.
 FT SIGNAL 1 22
 FT PROPEP 23 45
 FT CHAIN 46 433
 FT PROPEP 434 514
 FT ACT SITE 278 278
 FT CARBOHYD 460 460
 FT CARBOHYD 472 472
 FT CARBOHYD 5 5
 FT CONFLICT 12 12
 FT CONFLICT 34 35
 FT CONFLICT 37 37
 FT CONFLICT 88 88
 FT CONFLICT 98 98
 FT CONFLICT 451 451
 FT CONFLICT 454 454
 FT CONFLICT 504 504
 FT CONFLICT 507 507
 SQ SEQUENCE 514 AA; 56645 MW; 62461IC3FABD6302 CRC64;
 Query Match 100.0%; Score 76; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LGFTLMGKG 15
 DB 140 LGFTLMGKG 154

RESULT 2
 ID YE12 YEAST STANDARD; PRT; 231 AA.
 AC P39975;
 DT 01-PEB-1995 (Rel. 31, Created)
 DT 01-PEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 26.8 kDa protein in DMD3 5' region.
 GN YEL072W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=6288c / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Huntke-Smith S.,
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U18795; AAB65015.1; -
 DR SGD; S0000798; YEL072W.
 KM Hypothetical protein.
 SQ SEQUENCE 231 AA; 26615 MW; 1C6A1A3CE19D5B1C CRC64;
 Query Match 52.6%; Score 40; DB 1; Length 231;
 Best Local Similarity 63.6%; Pred. No. 5.3;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 LGFTLMGKG 15
 DB 137 LGFTLMGKG 147

RESULT 3
 ID SUH3_RAT STANDARD; PRT; 285 AA.
 AC P50235;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alcohol sulfotransferase (EC 2.8.2.2) (Hydroxysteroid
 DE sulfotransferase) (ST) (ST-60).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RA MEDLINE=94306585; Pubmed=8033273;
 RA Medabe T., Ogura K., Satsukawa M., Okuda H., Hiratsuka A.;
 RT "Molecular cloning and functions of rat liver hydroxysteroid
 RT sulfotransferases catalysing covalent binding of carcinogenic
 RT polycyclic arylmethanols to DNA."
 RL Chem. Biol. Interact. 92:87-105(1994).
 CC -1- FUNCTION: CATALYZES SULFATION OF HYDROXYSTERIODS AND XENOBIOTICS
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + an alcohol =
 CC adenosine 3',5'-bisphosphate + an alkyl sulfate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- INDUCTION: INDUCED BY ESTROGENS AND SUPPRESSED BY ANDROGENS.
 CC EXPRESSION IS UNDER THE INFLUENCE OF PITUITARY GROWTH HORMONE AND
 CC THYROID HORMONE.
 CC -1- SIMILARITY: BELONGS TO THE SULFOTRANSFERASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D14989; BAA03634.1; -
 DR HSSP; P50224; ICM.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KM Transferase; Steroid metabolism.
 FT BINDING 249 255
 SQ SEQUENCE 285 AA; 33531 MW; 5ACAS6A859F9D4D5 CRC64;
 Query Match 52.6%; Score 40; DB 1; Length 285;
 Best Local Similarity 80.0%; Pred. No. 6.7;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 6 LGFTLMGKG 15
 DB 240 LGFTLMGKG 249

RESULT 4
 ID PROB_THEMA STANDARD; PRT; 353 AA.
 AC Q9WYD0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glutamate 5-kinase (EC 2.7.2.11) (Gamma-glutamyl kinase) (GK).
 GN PROB OR TW0294.
 OS Thermotoga maritima.

```

OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OC Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS8 / DSM 3109;
RX MEDLINE=9287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Lither K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -1- FUNCTION: Catalyzes the transfer of a phosphate group to glutamate
CC to form glutamate 5-phosphate which rapidly cyclizes to 5-
CC oxoprolinone.
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate = ADP + L-glutamate 5-
CC phosphate.
CC -1- PATHWAY: Proline biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GLUTAMATE 5-KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001711; AAD35382.1; -
DR TIGR; TM0294; -
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR001057; Glu_5kinase.
DR InterPro; IPR002478; PUA_5kinase.
DR Pfam; PF00696; aak_kinase; 1.
DR Pfam; PF01472; PUA_1.
DR PRINTS; PR00474; GLU5KINASE.
DR SMART; SM00359; PUA; 1.
DR TIGRPFAM; TIGR01027; prob; 1.
DR PROSITE; PS00902; GLUTAMATE_5_KINASE; 1.
DR TRANSFERASE; Kinase; Prolinase biosynthesis; Complete proteome.
SQ SEQUENCE 353 AA; 38320 MW; 72E3171CB9F760 CRC64;

Query Match 52.6%; Score 40; DB 1; Length 353;
Best Local Similarity 58.3%; Pred. No. 8.3;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 AKLTGFTLMKG 15
|:|||||
DB 49 ARAAGFTYLGKG 60

RESULT 5
Y095_SINY3 STANDARD; PRT; 567 AA.
AC Q55884;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein s110095.
GN S110095.
OC Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;

```

```

RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64k to 92k of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- SIMILARITY: BELONGS TO THE ABC1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D64004; BAA10646.1; -
DR InterPro; IPR004147; ABC1.
DR Pfam; PF03109; ABC1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 567 AA; 64990 MW; 1D52C6A6934CA24 CRC64;

Query Match 52.6%; Score 40; DB 1; Length 567;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LQFAKLGTFLM 12
|:|||||
DB 508 LIFAKLTGFTLL 519

RESULT 6
P01_RABIT STANDARD; PRT; 358 AA.
ID P01_RABIT
AC P27170;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serum paroxonase/arylesterase 1 (EC 3.1.1.2) (EC 3.1.8.1) (PON 1)
DE (Serum arylalkylphosphatase 1) (A-esterase 1) (Aromatic esterase 1).
GN PON1 OR PON.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=92031445; PubMed=1657140;
RA Hasselet C., Richter R.J., Humbert R., Chapline C., Crabb J.W.,
RA Omiecinski C.J., Furlong C.E.;
RT "Characterization of cDNA clones encoding rabbit and human serum
RT paroxonase: the mature protein retains its signal sequence.";
RL Biochemistry 30:10141-10149(1991).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=93345100; PubMed=8393745;
RA Furlong C.E., Costa L.G., Hasselet C., Richter R.J.,
RA Sundstrom J.A., Adler D.A., DiStefano C.M., Omiecinski C.J.,
RA Chapline C., Crabb J.W.;
RT "Human and rabbit paroxonases: purification, cloning, sequencing,
RT mapping and role of polymorphism in organophosphate detoxification.";
RL Chem. Biol. Interact. 87:35-48(1993).
RN [3]
RP CHARACTERIZATION, AND SEQUENCE OF 1-20.
RX MEDLINE=92031444; PubMed=1718413;
RA Furlong C.E., Richter R.J., Chapline C., Crabb J.W.;
RT "Purification of rabbit and human serum paroxonase.";
RL Biochemistry 30:10133-10140(1991).
CC -1- FUNCTION: HYDROLYZES THE TOXIC METABOLITES OF A VARIETY OF
CC ORGANOPHOSPHORUS INSECTICIDES. CAPABLE OF HYDROLYZING A BROAD
CC SPECTRUM OF ORGANOPHOSPHATE SUBSTRATES AND A NUMBER OF AROMATIC
CC CARBOXYLIC ACID ESTERS.
CC -1- CATALYTIC ACTIVITY: Aryl dialkyl phosphate + H(2)O = dialkyl
CC phosphate + an aryl alcohol.

```


RA Yamamoto Y., Horiuchi T.;
 RT "A 570-Kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map."
 RL DNA Res. 3:363-377(1996).
 RN [4]
 RX CHARACTERIZATION AS AN HEAT-SHOCK GENE.
 RA MEDLINE=93352431; PubMed=8349564;
 RT Chuang S.B., Blatner F.R.;
 RT "Characterization of twenty-six new heat shock genes of Escherichia
 RT coli."
 RL J. Bacteriol. 175:5242-5252(1993).
 CC -1- FUNCTION: FERMENTATIVE LACTATE DEHYDROGENASE.
 CC -1- CATALYTIC ACTIVITY: (R)-lactate + NAD(+) = pyruvate + NADH.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U36928; AAB51772.1; -
 DR EMBL: AB000235; AAC74462.1; -
 DR EMBL: D90775; BA14985.1; -
 DR EMBL: D90776; BA14990.1; -
 DR HSSP: P30901; 2DLD.
 DR ECO2DBASE: D033.4; 6TH EDITION.
 DR EcGene: EG3166; 1dbA.
 DR InterPro: IPR002162; D_2hyd.ac.dh.
 DR Pfam: PF00389; 2-Hacid.DH.1.
 DR Pfam: PF02826; 2-Hacid.DH.C.1.
 DR PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.
 KM Oxidoreductase; NAD; Heat shock; Complete proteome.
 FT NP BIND 146 174 NAD (BY SIMILARITY).
 FT ACT SITE 234 263 SUBSTRATE-BINDING (BY SIMILARITY).
 FT ACT SITE 263 295 BY SIMILARITY.
 FT ACT SITE 295 295 BY SIMILARITY.
 SQ SEQUENCE 329 AA; 36534 MW; EF85419986438D6D CRC64;
 Query Match 50.0%; Score 38; DB 1; Length 329;
 Best Local Similarity 77.8%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 6 LTGFTLMGK 14
 DB 137 LTGFTLMGK 145
 RESULT 9
 DHYS_HUMAN STANDARD; PRT; 369 AA.
 ID P49366; O13276; Q33184;
 AC P49366; O13276; Q33184;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Deoxyhypusine synthase (EC 2.5.1.46) (DHS).
 GN DHS OR DS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OK NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95403436; PubMed=7673224;
 RA Joe Y.A., Wolfe E.C., Park M.H.;
 RT "Cloning and expression of human deoxyhypusine synthase cDNA.
 RT Structure-function studies with the recombinant enzyme and mutant
 RT proteins.";

RL J. Biol. Chem. 270:22386-22392(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=96140738; PubMed=8549832;
 RA Bevec D., Kappel B., Jakse H., Geonga R., Hauber J., Klier H.,
 RA Steinkasserer A.;
 RT "Molecular characterization of a cDNA encoding functional human
 RT deoxyhypusine synthase and chromosomal mapping of the corresponding
 RT gene locus."
 RL FEBS Lett. 378:195-198(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96201991; PubMed=8615810;
 RA Yan Y.P., Tao Y., Chen K.Y.;
 RT "Molecular cloning and functional expression of human deoxyhypusine
 RT synthase cDNA based on expressed sequence tag information."
 RL Biochem. J. 315:429-434(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97264341; PubMed=9110174;
 RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
 RA Ricchetti J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
 RT "Large-scale concatenation cDNA sequencing."
 RL Genome Res. 7:353-358(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98332729; PubMed=9666110;
 RA Mantuano E., Trettel F., Olsen A.S., Lemm G., Frontali M.,
 RA Jodice C.;
 RT "Localization and genomic structure of human deoxyhypusine synthase
 RT gene on chromosome 19p13.2-distal 19p13.1."
 RL Gene 215:153-157(1998).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain; and Lung;
 RA Strausberg R.;
 RN Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=98154315; PubMed=9493264;
 RA Liao D.-I., Wolff E.C., Park M.H., Davies D.R.;
 RT "Crystal structure of the NAD complex of human deoxyhypusine
 RT synthase: an enzyme with a ball-and-chain mechanism for blocking the
 RT active site."
 RL Structure 6:23-32(1998).
 CC -1- FUNCTION: CATALYZES THE NAD-DEPENDENT OXIDATIVE CLEAVAGE OF
 CC SPERMIDINE AND THE SUBSEQUENT TRANSFER OF THE BUTYLAMINE MOIETY OF
 CC SPERMIDINE TO THE EPSILON-AMINO GROUP OF A SPECIFIC LYSINE RESIDUE
 CC OF THE EIF-5A PRECURSOR PROTEIN TO FORM THE INTERMEDIATE
 CC DEOXYHYPUSINE RESIDUE.
 CC -1- CATALYTIC ACTIVITY: [eIF5A-precursor]-lysine + spermidine =
 CC [eIF5A-precursor]-deoxyhypusine + propan-1,3-diamine.
 CC -1- COFACTOR: NAD.
 CC -1- PATHWAY: Hypusine biosynthesis; first step.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. THE SHORT FORM
 CC IS INACTIVE.
 CC -1- SIMILARITY: BELONGS TO THE DEOXYHYPUSINE SYNTHASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L39068; AAA86282.1; -
 DR EMBL: U40579; AAA96151.1; -
 DR EMBL: U32178; AAB02179.1; -
 DR EMBL: U26266; AAB02175.1; -

DR EMBL; U79262; AAB50208.1; --
 DR EMBL; AJ001701; CA04940.1; --
 DR EMBL; AJ001702; CA04940.1; JOINED.
 DR EMBL; AJ001703; CA04940.1; JOINED.
 DR EMBL; AJ001704; CA04940.1; JOINED.
 DR EMBL; BC000333; AAH0033.1; --
 DR EMBL; BC014016; AAH14016.1; --
 DR PDB; 1DHS; 25-FEB-98.
 DR Genew; HGNC:2869; DHS.
 DR MIM; 600944; --
 DR InterPro; IPR002773; DS.
 DR Pfam; PF01916; DS; 1.
 DR ProDom; PD007730; DS; 1.
 DR TIGRFAMs; TIGR00321; dhys; 1.
 KM Hypusine biosynthesis; Transferase; NAD; Alternative splicing;
 3d-structure.
 FT VARSPIC 262 308 MISSING (IN SHORT ISOCORM).
 FT CONFLICT 11 11 A -> R (IN REF. 2 AND 3).
 FT CONFLICT 13 14 AL -> R (IN REF. 3).
 FT CONFLICT 85 85 A -> G (IN REF. 3).
 FT CONFLICT 196 196 T -> I (IN REF. 3).
 FT CONFLICT 199 199 V -> A (IN REF. 3).
 FT CONFLICT 220 220 V -> A (IN REF. 3).
 FT CONFLICT 296 297 MR -> SG (IN REF. 3).
 FT CONFLICT 311 311 E -> EE (IN REF. 3).
 SQ SEQUENCE 369 AA; 40970 MW; 5314FED620AC9BE7 CRC64;

Query Match 50.0%; Score 38; DB 1; Length 369;
 Best Local Similarity 53.8%; Pred. No. 20;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 FAKLTGFTLMGKG 15
 Db 272 FAKLTGFTLMGKG 284

RESULT 10
 RRP3_IATIE STANDARD; PRT; 759 AA.
 ID RRP3_IATIE
 AC P18884;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE RNA-directed RNA polymerase subunit p3 (EC 2.7.7.48) (Polymerase basic
 protein 2) (PB2).
 OS Influenza A virus (strain A/Kiev/59/79).
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 CC Influenza A viruses; Influenzavirus A.
 NCBI_Taxid=11422;
 RX MEDLINE=8049805; PubMed=3675640;
 RA Petrov N.A., Netesov S.V., Golovin S.Y., Mamaeva N.V.,
 RA Mamaev L.V., Sivododova G.F., Petrenko V.A., Vasilenko S.K.;
 RT "Primary structure of the full-length DNA copy of the influenza virus
 A/Kiev/59/79 (H1N1) PB2 gene";
 RL Bioorg. Khim. 13:915-920(1987).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 (RNA) (N).
 CC -1- SUBUNIT: INFLUENZA RNA POLYMERASE IS COMPOSED OF THREE SUBUNITS:
 P1 (OR PB1), P2 (OR PA), AND P3 (OR PB2).
 CC -1- SIMILARITY: BELONGS TO THE INFLUENZA VIRUSES POLYMERASE PB2
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M38277; AAA43653.1; --

DR PIR; J00411; P31YAK.
 DR InterPro; IPR001591; Flu_PB2.
 DR Pfam; PF00604; Flu_PB2; 1.
 DR ProDom; PD001667; Flu_PB2; 1.
 KM Transferase; RNA-directed RNA polymerase.
 SQ SEQUENCE 759 AA; 86000 MW; 9272580CEAF05AFC CRC64;

Query Match 50.0%; Score 38; DB 1; Length 759;
 Best Local Similarity 46.7%; Pred. No. 43;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LOPAKLTGFTLMGKG 15
 Db 686 VESAVLRGFTLMGKG 700

RESULT 11
 YMF9_YEAST STANDARD; PRT; 1679 AA.
 ID YMF9_YEAST
 AC 004358;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 187.1 kDa protein in OG31-CNA2 intergenic region.
 GN YML059C OR YMF958.03C.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_Taxid=4932;
 RX STRAIN=S288c / AB972;
 RA Devlin K., Churcher C., Barrell B.G., Rajandream M.A.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE YMF028 (SMS) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; Z46729; CA86716.1; --
 DR SGD; S0004524; YML059C.
 DR InterPro; IPR002641; Patatin.
 DR InterPro; IPR001423; YPF0028.
 DR InterPro; IPR000595; CNMP_binding.
 DR Pfam; PF00027; CNMP_binding; 2.
 DR Pfam; PF01734; Patatin; 1.
 DR SMART; SM00100; CNMP; 1.
 DR PROSITE; PSS0042; CNMP_BINDING_3; 2.
 DR PROSITE; PSS01237; YPF0028; 1.
 DR Hypothetical protein; Transmembrane.
 FT TRANSMEM 50 70 POTENTIAL.
 FT TRANSMEM 104 124 POTENTIAL.
 SQ SEQUENCE 1679 AA; 187132 MW; 550FFCD4ACAF8B25 CRC64;

Query Match 49.3%; Score 37.5; DB 1; Length 1679;
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;
 Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 LOPAKLTGFTLMGKG 15
 Db 1280 IQFS-LITFTLMGKG 1293

RESULT 12
 SPAP_SALT1 STANDARD; PRT; 224 AA.
 ID SPAP_SALT1
 AC Q56023;
 DT 01-NOV-1997 (Rel. 35, Created)

ID CYBH ALCEU STANDARD; PRT; 244 AA.
 AC P31898;
 DT 01-JUN-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Probable Ni/Fe-hydrogenase B-type cytochrome subunit.
 GN HOXZ.
 OS Alkaligenes eutrophus (Ralstonia eutropha).
 OC Plasmid megaplasmid pHG1.
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=510;
 RN SEQUENCE FROM N.A.
 RC STRAIN=H16 / DSM 428 / ATCC 17699;
 RX MEDLINE=93015670; PubMed=1383192;
 RA Kortliene C., Horstmann K., Schwartz E., Rohde W., Binsack R.,
 RA Friedrich B.;
 RT "A gene complex coding for the membrane-bound hydrogenase of
 RT Alkaligenes eutrophus H16.";
 RL J. Bacteriol. 174:6277-6289(1992).
 CC -1- FUNCTION: PROBABLE B-TYPE CYTOCHROME.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE HDPC/HYAC/HYDC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M96433; AAA16463.1; -.
 DR PIR: C43255; C43255.
 DR InterPro: IPR000516; Ni_hydr_Cytr.
 DR Pfam: PF01292; Ni_hydr_Cytr_1.
 DR PRINTS: PR00161; N1HGENSECTB.
 DR PRODOM: PD005374; N1_hydr_Cytr_1.
 DR PROSITE: PS00882; N1_HGENSE CTB_1; 1.
 DR PROSITE: PS00883; N1_HGENSE CTB_2; 1.
 KW Transmembrane; Electron transport; Heme; Plasmid.
 FT TRANSMEM 39 59 POTENTIAL.
 FT TRANSMEM 73 93 POTENTIAL.
 FT TRANSMEM 150 171 POTENTIAL.
 FT TRANSMEM 204 221 POTENTIAL.
 SQ SEQUENCE 244 AA; 27581 MW; 49435C5ECC706961 CRC64;
 Query Match 48.7%; Score 37; DB 1; Length 244;
 Best Local Similarity 53.8%; Pred. No. 20; Mismatches 4; Indels 0; Gaps 0;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 3 FAKLTGFTLMGKG 15
 DB 162 FMILTGPAMYGEG 174
 RESULT 15
 ID LPXA CHLMU STANDARD; PRT; 280 AA.
 AC 09PUL1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acyl-[acyl-carrier-protein]-UDP-N-acetylglucosamine O-acyltransferase
 DE (EC 2.3.1.129) (UDP-N-acetylglucosamine acyltransferase).
 GN LPXA OR TC0818.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83560;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=MOPn / N199;

RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uetendack T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gaim M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 CC -1- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF LIPID A, A
 CC PHOSPHORYLATED GLYCOLIPID THAT ANCHORS THE LIPOLYSACCHARIDE TO
 CC THE OUTER MEMBRANE OF THE CELL. (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: (R)-3-hydroxytetradecanoyl-[acyl-carrier
 CC protein] + UDP-N-acetylglucosamine = [acyl-carrier protein] +
 CC UDP-3-O-(3-hydroxytetradecanoyl)-N-acetylglucosamine.
 CC -1- PATHWAY: Lipid A biosynthesis; first step.
 CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERASE HEXAPEPTIDE REPEAT FAMILY.
 CC LPXA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE002347; AAF39620.1; -.
 DR HSSP: P10440; 1LXA.
 DR TIGR: TC0818; -.
 DR InterPro: IPR001451; Hexapep_transf.
 DR Pfam: PF00132; hexapep_8.
 DR PROSITE: PS00101; HEXAPEP_TRANSFERRASES; 1.
 KW Transferase; Acyltransferase; Lipid A biosynthesis; Lipid synthesis;
 KW Repeat; Complete proteome.
 SQ SEQUENCE 280 AA; 30672 MW; 4EE1B58937F0A63F CRC64;
 Query Match 48.7%; Score 37; DB 1; Length 280;
 Best Local Similarity 46.2%; Pred. No. 23; Mismatches 3; Indels 0; Gaps 0;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 OY 3 FAKLTGFTLMGKG 15
 DB 43 YAVIDGFTTIGRG 55

Search completed: April 20, 2003, 13:07:41
 Job time : 3.92105 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 12.6711 seconds
(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524d-101

Perfect score: 76

Sequence: 1 LQFAKLTGFTLMGK 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PLANT:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	63.2	507	10	Q9FY19 juniperus a
2	45	59.2	474	10	Q8W4S2 arbidopsi
3	45	59.2	540	10	Q226S9 arbidopsi
4	44	57.9	475	10	Q94EH4 arbidopsi
5	44	57.9	475	10	Q949Z1 arbidopsi
6	44	56.6	482	10	Q9LNG3 arbidopsi
7	43	56.6	194	16	Q9X7S4 streptomyc
8	43	56.6	439	10	Q9SWS2 streptomyc
9	43	56.6	443	10	Q9SWS3 streptomyc
10	43	56.6	457	10	Q9ZUE7 streptomyc
11	43	56.6	502	16	Q9RS41 streptomyc
12	43	56.6	519	10	Q9LPS6 arbidopsi
13	43	56.6	1161	10	Q9LQD1 arbidopsi
14	41	53.9	360	16	Q9KZK6 streptomyc
15	41	53.9	442	10	Q9SDP3 allium cepa
16	41	53.9	468	10	Q9CAL5 arbidopsi

17	41	53.9	887	16	Q8XY49 ralstonia s
18	40	52.6	217	16	Q8YGS0 bruceella me
19	40	52.6	355	10	Q94KI6 pisum sativ
20	40	52.6	358	17	Q9YBY3 aeropyrum p
21	40	52.6	371	17	Q28833 archaeoglob
22	40	52.6	383	6	Q95IV2 macaca fasc
23	40	52.6	435	10	Q9LYJ5 arbidopsi
24	40	52.6	578	6	Q95JY8 macaca fasc
25	40	52.6	743	6	Q8WPJ3 macaca fasc
26	39	51.3	170	2	Q9X3N0 neisseria m
27	39	51.3	193	16	Q9CMN2 pasteurella
28	39	51.3	217	16	Q8YFT0 bruceella me
29	39	51.3	345	10	Q94CZ5 oryza sativ
30	39	51.3	359	6	Q9BGN2 oryctolagus
31	39	51.3	359	6	Q9BGN1 oryctolagus
32	39	51.3	380	10	Q94BV8 arbidopsi
33	39	51.3	382	10	Q95IM7 arbidopsi
34	39	51.3	425	10	Q8RYE4 arbidopsi
35	39	51.3	519	2	Q9RCI4 oryza sativ
36	39	51.3	542	10	Q9SGY5 bacteroides
37	39	51.3	572	16	Q9YTM8 neisseria m
38	39	51.3	572	16	Q9JTM7 neisseria m
39	39	51.3	591	16	Q9J3L0 pseudomonas
40	39	51.3	689	10	Q9FPH8 arbidopsi
41	39	51.3	1020	16	Q9PB07 xylella fas
42	39	51.3	1270	5	Q9GPN0 caenorhabdi
43	39	51.3	1348	5	Q9VDT3 drosophila
44	38	50.0	244	2	Q9ANQ8 bradyrhizob
45	38	50.0	312	10	Q9SUV3 arbidopsi

ALIGNMENTS

RESULT 1	Q9FY19	PRELIMINARY;	PRT;	507 AA.
ID	Q9FY19	01-MAR-2001 (TREMBLrel. 16, Created)		
AC	Q9FY19	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Pollen major allergen 2 protein precursor.			
GN	JNA2.			
OS	Juniperus ashei (Ozark white cedar).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.			
OX	NCBI_TaxId=13101;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=MALE POLLEN;			
RX	MEDLINE=20403896; PubMed=10944464;			
RA	Yokoyama M., Miyahara M., Shimizu K., Kino K., Tsunoo H.;			
RT	"Purification, Identification and cDNA cloning of Jun a 2, the second			
RT	major allergen of mountain cedar pollen."			
RL	Biochem. Biophys. Res. Commun. 275:195-202(2000).			
CC	-1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES			
CC	(POLYGLACTURONASES).			
CC	EMBL: AJ404653; CAC0582.1; -.			
DR	HSSP; P26509; IHR.			
DR	InterPro; IPR000743; GH28.			
DR	InterPro; IPR000408; Reg_chir condens.			
DR	Pfam; PF00295; Glyco_hydro_28; 1.			
DR	PROSITE; PS00502; POLYGLACTURONASE; UNKNOWN_1.			
DR	PROSITE; PS00626; RC01_2; UNKNOWN_1.			
KW	Cell wall, Glycosidase, Hydrolase, Signal.			
FT	SIGNAL 1			
FT	POTENTIAL.			
FT	SEQUENCE 507 AA; 55730 MW; 282EOA5B958FE5A CRC64;			
QY	Query Match	63.2%; Score 48; DB 10; Length 507;		
QY	Best Local Similarity	60.0%; Pred. No. 3.1;		
QY	Matches	9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;		
QY	1 LQFAKLTGFTLMGK 15			

DB 141 MHPALTDENLMTGTG 155

RESULT 2

Q8W4S2 PRELIMINARY; PRT; 474 AA.
 ID Q8W4S2
 AC Q8W4S2;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE At1g60590/F8A5_12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY060568; AAL31197.1; -.
 DR InterPro; IPR000743; GH28.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.
 DR SEQUENCE 474 AA; 51964 MW; 9631844B83E209C6 CRC64;
 SQ
 Query Match 59.2%; Score 45; DB 10; Length 474;
 Best Local Similarity 46.7%; Pred. No. 9.9; Indels 0; Gaps 0;
 Matches 7; Conservative 4; Mismatches 4;
 QY 1 LQFAKLTGFTLMGKG 15
 DB 140 IEFKLSGITIQNG 154
 RESULT 3
 ID Q22699 PRELIMINARY; PRT; 540 AA.
 AC Q22699;
 DT 01-JUN-1998 (TREMBlrel. 05, Created)
 DT 01-JUN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative polygalacturonase.
 GN F8A5_12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Federespiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
 RA Lu M., Araujo R., Brendel V., Buehler E., Dewar K., Feng J., Kim C.,
 RA Li Y., Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M.,
 RA Vysotskaya V.S., Yu G., Ecker J., Theologis A., Davis R.W.;
 RT "Genomic sequence of Arabidopsis."
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 DE EMBL; AC002292; AAB71972.1; -.
 DR InterPro; IPR000743; GH28.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 CX NCBI_TaxID=3702;
 KM Cell wall; Glycosidase; Hydrolase.

SQ SEQUENCE 540 AA; 59631 MW; 532B79A84EBB904 CRC64;

Query Match 59.2%; Score 45; DB 10; Length 540;
 Best Local Similarity 46.7%; Pred. No. 11;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGKG 15

DB 206 IEFKLSGITIQNG 220

RESULT 4

Q94EH4 PRELIMINARY; PRT; 475 AA.
 ID Q94EH4
 AC Q94EH4;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE At1g48100/F2ID18_17.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 DE EMBL; AF410319; AAK95305.1; -.
 DR InterPro; IPR000743; GH28.
 DR InterPro; IPR000408; Reg_chromatins.
 DR InterPro; IPR001412; RNA-synt_1.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_1; UNKNOWN_1.
 DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.
 DR PROSITE; PS00626; RCCL_2; UNKNOWN_1.
 KM Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 475 AA; 51380 MW; 4A42C4355827D30A CRC64;
 Query Match 57.9%; Score 44; DB 10; Length 475;
 Best Local Similarity 60.0%; Pred. No. 15;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 LQFAKLTGFTLMGKG 15
 DB 164 LVFRLDGFPSGKG 178
 RESULT 5
 ID Q949Z1 PRELIMINARY; PRT; 475 AA.
 AC Q949Z1;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative polygalacturonase PGL.
 GN F2ID18_18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 CX NCBI_TaxID=3702;
 KM

RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
 RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
 RA Carinini P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamuya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B.,
 RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shimizu P., Southwick A., Tracy S.E.,
 RA Shinzaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT "Full length cDNA of gene F1D18.18 (GI:8778527)."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYACTURONASES).
 DR EMBL; AY050798; AAK92733.1; -.
 DR InterPro; IPR000743; GH28.
 DR InterPro; IPR000408; Reg. chr. condens.
 DR InterPro; IPR001412; tRNA-synt.1.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_1; UNKNOWN_1.
 DR PROSITE; PS00502; POLYACTURONASE; UNKNOWN_1.
 DR Cell wall; Glycosidase; RCC1_2; UNKNOWN_1.
 KM SEQUENCE 475 AA; 51352 MW; ACB066CF725F2C8F CRC64;
 SQ

Query Match 57.9%; Score 44; DB 10; Length 475;
 Best Local Similarity 60.0%; Pred. No. 15;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGKG 15
 DB 164 LVFRLDGFTEGSGK 178

RESULT 6

ID Q9LNG3 PRELIMINARY; PRT; 492 AA.
 AC Q9LNG3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE F21D18.18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shim P., Altafi H., Bei Q., Chin C., Chou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F21D18 from chromosome
 1".
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chou J., Choi E.,
 RA Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu S., Mukharzky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,

RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYACTURONASES).
 DR EMBL; AC023673; AAF79535.1; -.
 DR InterPro; IPR000743; GH28.
 DR InterPro; IPR000408; Reg. chr. condens.
 DR InterPro; IPR001412; tRNA-synt.1.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_1; UNKNOWN_1.
 DR PROSITE; PS00502; POLYACTURONASE; 1.
 DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
 DR Cell wall; Glycosidase; Hydrolase.
 KM SEQUENCE 492 AA; 53395 MW; 42BAF98B9E30AC1A CRC64;
 SQ

Query Match 57.9%; Score 44; DB 10; Length 492;
 Best Local Similarity 60.0%; Pred. No. 16;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGKG 15
 DB 164 LVFRLDGFTEGSGK 178

RESULT 7

ID Q9X7S4 PRELIMINARY; PRT; 194 AA.
 AC Q9X7S4;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein SC07276.
 GN SC07276 OR SC5H1.16.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Kieeer H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornby T., Howarth S.,
 RA Huang C.-H., Kieeer T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)".
 RL Nature 417:141-147(2002).
 DR EMBL; AL049863; CAB42941.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 194 AA; 20485 MW; D4BR04D35971F875 CRC64;

Query Match 56.6%; Score 43; DB 16; Length 194;
 Best Local Similarity 63.6%; Pred. No. 9.3;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 KLTGFTLMGKG 15
 DB 37 RITGFTLMGKG 47

RESULT 8

ID Q9SWS2 PRELIMINARY; PRT; 439 AA.
 AC Q9SWS2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Polyalacturonase PG2.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99285288; PubMed=10356800;
 RA Mahalingam R., Wang G., Knap H.T.;
 RT "Polyalacturonase and polyalacturonase inhibitor protein: gene
 isolation and transcription in Glycine max-Heterodera glycines
 RT interaction.";
 RL Mol. Plant Microbe Interact. 12:490-498 (1999).
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 DR EMBL: AF128267; AAD46484.1; -;
 DR InterPro: IPR000743; GH28.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 KW Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 439 AA; 47018 MW; D07AA96E3F42B7DB CRC64;

Query Match 56.6%; Score 43; DB 10; Length 439;
 Best Local Similarity 53.3%; Pred. No. 21;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 LOFAGLTGFTLMGKG 15
 Db 125 LEFSKLNKRTIRGKG 139

RESULT 9
 O9SMS3 PRELIMINARY; PRT; 443 AA.
 AC O9SMS3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Polyalacturonase PG1.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99285288; PubMed=10356800;
 RA Mahalingam R., Wang G., Knap H.T.;
 RT "Polyalacturonase and polyalacturonase inhibitor protein: gene
 RT isolation and transcription in Glycine max-Heterodera glycines
 RT interactions.";
 RL Mol. Plant Microbe Interact. 12:490-498 (1999).
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 DR EMBL: AF128267; AAD46483.1; -;
 DR InterPro: IPR000743; GH28.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 KW Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 443 AA; 47335 MW; 13314187F6EBF33 CRC64;

Query Match 56.6%; Score 43; DB 10; Length 443;
 Best Local Similarity 53.3%; Pred. No. 21;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 LOFAGLTGFTLMGKG 15
 Db 129 LEFSKLNKRTIRGKG 143

RESULT 10
 O9ZUE7

ID O9ZUE7 PRELIMINARY; PRT; 457 AA.
 AC O9ZUE7;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE F508.2 protein.
 GN F508.2
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Liu S., Lenz C.,
 RA Li J., Kremenetskaia I., Luros J., Altfelt H., Gonzalez A., Araujo R.,
 RA Buehler E., Conn L., Conway A.B., Dunn P., Hansen N., Hultzer L.,
 RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
 RA Ecker J.R., Federiguel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F508 sequence.";
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 DR EMBL: AC005990; AAC98004.1; -;
 DR InterPro: IPR000886; ER_target.
 DR InterPro: IPR000743; GH28.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 DR PROSITE: PS00014; ER_TARGET; UNKNOWN 1.
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 KW Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 457 AA; 49269 MW; OEAA6551E566168 CRC64;

Query Match 56.6%; Score 43; DB 10; Length 457;
 Best Local Similarity 53.3%; Pred. No. 22;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 LOFAGLTGFTLMGKG 15
 Db 154 LEFSKLNKGVPOGKG 168

RESULT 11
 O9RS41 PRELIMINARY; PRT; 502 AA.
 AC O9RS41;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein DR2286.
 GN DR2286.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R1;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson K.J., Hatt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Vang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 radiodurans R1.";

DR EMBL: AC007945; AAF79584.1; -
 DR InterPro: IPR000886; ER target.
 DR InterPro: IPR000743; CH28.
 DR InterPro: IPR002885; PPR.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 DR Pfam: PF01535; PPR; 11.
 DR TIGRFAMs: TIGR00756; PPR; 9.
 DR PROSITE: PS00014; ER_TARGET; UNKNOWN 1.
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 DR Cell wall; Glycosidase; Hydrolase.
 KM SEQUENCE: 1161 AA; 128197 MW; C67DCFEBCBCAB9 CRC64;
 SQ SEQUENCE: 1161 AA; 128197 MW; C67DCFEBCBCAB9 CRC64;
 Query Match 56.6%; Score 43; DB 10; Length 1161;
 Best Local Similarity 53.3%; Pred. No. 56;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 LQPAKLTGFTLMGK 15
 DB 827 LEFSKLGWVFOGK 841
 RESULT 14
 ID Q9KZK6 PRELIMINARY; PRT; 360 AA.
 AC Q9KZK6;
 DT 01-OCT-2000 (TREMBlrel. 15; Created)
 DT 01-OCT-2000 (TREMBlrel. 15; Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21; Last annotation update)
 DE Putative nucleotide phosphorylase.
 GN SC03039 OR SCE34.20C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Saunders D.C.; Harris D.;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=843436;
 RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
 RA Kinsahl H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 DR EMBL: AL353862; CAB8923.1; -
 DR InterPro: IPR001451; Hexapep_transf.
 DR InterPro: IPR001825; NTP_transferase.
 DR Pfam: PF00132; hexaped; 3.
 DR Pfam: PF00483; NTP_transferase; 1.

SQ SEQUENCE 360 AA; 37642 MW; 9C50CDF2B2CCE515 CRC64;
 Query Match 53.9%; Score 41; DB 16; Length 360;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 4 AKLTGFTLMGK 15
 DB 266 AKLTGFTVBSG 277
 RESULT 15
 ID Q9SDP3 PRELIMINARY; PRT; 442 AA.
 AC Q9SDP3;
 DT 01-MAY-2000 (TREMBlrel. 13; Created)
 DT 01-MAY-2000 (TREMBlrel. 13; Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21; Last annotation update)
 DE APS-reductase.
 OS Allium cepa (Onion).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
 OC Allium.
 OX NCBI_TaxID=4679;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. PUKECOHE LONG KEEPER; TISSUE=ROOT;
 RA Pither-Joyce M.D., McCallum J.A.;
 RT "Isolation of an APS-reductase cDNA clone from Allium cepa.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF212155; AAF18999.1; -
 DR InterPro: IPR004508; APS_reduc.
 DR InterPro: IPR002500; PAPS_reduct.
 DR InterPro: IPR00063; Thioired.
 DR Pfam: PF01507; PAPS_reduct; 1.
 DR Pfam: PF00085; thioired; 1.
 DR TIGRFAMs: TIGR00424; APS_reduc; 1.
 SQ SEQUENCE 442 AA; 48710 MW; BE570C4F84A490CF CRC64;
 Query Match 53.9%; Score 41; DB 10; Length 442;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LQPAKLTGFTLMGK 14
 DB 127 IEXARLTGFDLVEK 140

Search completed: April 20, 2003, 13:13:14
 Job time : 14.6711 secs

GenCore version 5.1.4_p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds

(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524D-120

Sequence: 1 GIDIFASKNFHLQKN 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database:

1: PIR.73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	514	2	S48730
2	79	100.0	514	2	JC2498
3	62	78.5	514	2	JC7100
4	54	68.4	507	2	JC7366
5	43	54.4	419	2	T38199
6	42	53.2	305	2	T20885
7	42	53.2	305	2	T20885
8	41	51.9	269	2	T39029
9	41	51.9	269	2	T39029
10	41	51.9	269	2	T39029
11	40	50.6	1021	2	S64506
12	40	50.6	1021	2	S64506
13	39	49.4	280	2	T43075
14	39	49.4	280	2	T43075
15	39	49.4	280	2	T43075
16	39	49.4	280	2	T43075
17	39	49.4	280	2	T43075
18	39	49.4	280	2	T43075
19	39	49.4	280	2	T43075
20	39	49.4	280	2	T43075
21	38	48.1	161	2	T50639
22	38	48.1	161	2	T50639
23	38	48.1	161	2	T50639
24	38	48.1	161	2	T50639
25	38	48.1	161	2	T50639
26	38	48.1	161	2	T50639
27	38	48.1	161	2	T50639
28	38	48.1	161	2	T50639
29	38	48.1	161	2	T50639

ALIGNMENTS

30 37.5 47.5 1804 2 H96597
31 37 46.8 155 2 A31278
32 37 46.8 156 2 H97334
33 37 46.8 168 2 T08831
34 37 46.8 182 2 T46387
35 37 46.8 223 2 C86816
36 37 46.8 352 2 JU0153
37 37 46.8 353 2 T42971
38 37 46.8 509 1 DEH01P
39 37 46.8 509 1 UC4241
40 37 46.8 517 2 S65305
41 37 46.8 541 2 A81330
42 37 46.8 701 2 C84972
43 37 46.8 706 2 D90124
44 37 46.8 874 2 A10281
45 37 46.8 1031 2 T38411

hypothetical prote
interleukin-2 prec
hypothetical prote
disease resistance
hypothetical prote
hypothetical prote
metalloproteinase
hypothetical prote
dihydroliponamide d
probable membrane
probable glutathio
hypothetical prote
hypothetical prote
probable insectici
probable Grpase ac

RESULT 1
S48730
Cry j II protein - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C:Accession: S48730
R:Namba, M.; Kurose, M.; Torisoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.;
FEBS Lett. 353, 124-128, 1994
A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar
A:Reference number: S48730; MUID:95010777; PMID:7926035
A:Accession: S48730
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-514 <NAM>
A:Cross-references: GB:D37765; NID:9577695; PIDN:BA07021.1; PID:di007598; PID:9577694

Query Match 100.0%; Score 79; DB 2; Length 514;
Best Local Similarity 100.0%; Pred No. 2.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHLQKN 15
DB 235 GIDIFASKNFHLQKN 249

RESULT 2
JC2498
second major allergen Cry j II precursor - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C:Accession: JC2498; PC2346; A60147
R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese cedar
A:Reference number: JC2498; MUID:94271186; PMID:8002972
A:Accession: JC2498
A:Molecule type: mRNA
A:Residues: 1-514 <KOM>
A:Cross-references: DDBJ:D29772; NID:9506857; PIDN:BA06172.1; PID:9506858
A:Accession: PC2346
A:Molecule type: protein
A:Residues: 52-61 <KO2>
R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matunashi, T.
Allergy 45, 309-312, 1990
A:Title: Identification of the second major allergen of Japanese cedar pollen.
A:Reference number: A60147; MUID:90342988; PMID:2382797
A:Accession: A60147
A:Molecule type: protein
A:Residues: 55-64 <SAK>
C:Keywords: glycoprotein; pollen
F:1-54/Domain: signal sequence #status predicted <SIG>
F:55-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 79; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHLQKN 15
|||
Db 235 GIDIFASKNFHLQKN 249

RESULT 3

UC7100 polygalacturonase Cha o 2 - Japanese cypress

C/Species: Chaamecyparis obtusa (Japanese cypress)

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000

C/Accession: JG7100; PC7026

R/Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.

Biochem. Biophys. Res. Commun. 263, 166-171, 1999

A/Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all

A/Reference number: JG7100; PMID:10486272

A/Accession: JG7100

A/Molecule type: mRNA

A/Residues: 1-514 <MOR>

A/Accession: PC7026

A/Molecule type: protein

A/Residues: 51-62 <MO2>

Query Match 78.5%; Score 62; DB 2; Length 514;
Best Local Similarity 78.6%; Pred. No. 0.0035;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHLQKN 14
|||
Db 235 GIDIFASKNFHLQKN 248

RESULT 4

JG7366 Jun a 2 protein - mountain cedar

C/Species: Juniperus ashei (mountain cedar)

C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000

C/Accession: JG7366; PC7093

R/Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.

Biochem. Biophys. Res. Commun. 275, 195-202, 2000

A/Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all

A/Reference number: JG7366

A/Accession: JG7366

A/Molecule type: mRNA

A/Residues: 1-507 <YOK>

A/Accession: PC7093

A/Cross-references: GB:AJ04653

A/Molecule type: protein

A/Residues: 55-63 <YO2>

C/Comment: This protein, a second major allergen of mountain cedar pollen, which is inv

C/Keywords: glycoprotein; pollen

Query Match 68.4%; Score 54; DB 2; Length 507;
Best Local Similarity 71.4%; Pred. No. 0.097;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHLQKN 14
|||
Db 236 GIDIFASKNFHLQKN 249

RESULT 5

T38199 conserved hypothetical protein SPAC22F8.09 - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C/Accession: T38199

R.Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.; Ramsperger, U.; Pohl, T.
submitted to the EMBL Data Library, August 1999

A/Reference number: Z21777

A/Accession: T38199

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-419 <LYN>

A/Cross-references: EMBL:AL109831; PIDN:CA852719.1; GSPDB:GN00066; SPDB:SPAC22F8.09

A/Experimental source: strain 972h-; cosmid c22F8

C/Genetics:

A/Map position: 1

A/Introns: 46/2

Query Match 54.4%; Score 43; DB 2; Length 419;
Best Local Similarity 63.6%; Pred. No. 7.6;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 FASKNFHLQKN 15
|||
Db 407 YSHKDFHLQKN 417

RESULT 6

D82888 hypothetical protein UV459 [imported] - Ureaplasma urealyticum

C/Species: Ureaplasma urealyticum

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C/Accession: D82888

R/Glass, J.L.; Leikowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to Genbank, February 2000

A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a

A/Reference number: A82870

A/Accession: D82888

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-156 <GLA>

A/Cross-references: GB:AE002142; GB:AF222894; NID:G6899447; PIDN:AAF30871.1; GSPDB:GN

A/Experimental source: serovar 3; biovar 1

C/Genetics:

A/Map position: 5

A/Introns: 52/3; 82/1

Query Match 53.2%; Score 42; DB 2; Length 156;
Best Local Similarity 72.7%; Pred. No. 3.9;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 IDIFASKNFHL 12
|||
Db 66 IKIFKTNFHL 76

RESULT 7

T20685 hypothetical protein F10A3.12 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T20685

R/Lloyd, C.

submitted to the EMBL Data Library, March 1997

A/Reference number: Z19309

A/Accession: T20685

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-305 <WLL>

A/Cross-references: EMBL:Z92829; PIDN:CA807348.1; GSPDB:GN00023; CESP:F10A3.12

A/Experimental source: clone F10A3

C/Genetics:

A/Map position: 5

A/Introns: 52/3; 82/1

Query Match 53.2%; Score 42; DB 2; Length 305;

Best Local Similarity 72.7%; Pred. No. 8.2;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 IDIFASKNFHL 12
DB 112 IEIFAPKRFHL 122

RESULT 8

Q0BEP3

gene 58 protein - saimiriine herpesvirus 1 (strain 11)

C/Species: saimiriine herpesvirus 1

A/Note: host Saimiri sciureus (common squirrel monkey)

C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999

C/Accession: B36812

R/Albrecht, J.

submitted to the EMBL Data Library, January 1992

A/Description: Primary structure of the herpesvirus saimiri genome.

A/Reference number: A36806

A/Accession: B36812

A/Molecule type: DNA

A/Residues: 1-357 <ALB>

A/Cross-references: GB:X64346; NID:960320; PIDN:CAA45681.1; PID:960379

R/Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; W

A/Title: Primary structure of the herpesvirus saimiri genome.

A/Reference number: A37309; MUID:9233688; PMID:1321287

A/Contents: annotation; possible protein-coding frames

A/Note: neither amino acid nor nucleotide sequence is given

C/Genetics:

A/Gene: 58

C/Superfamily: human herpesvirus 4 BMR2 protein

Query Match

Best Local Similarity 53.2%; Score 42; DB 1; Length 357;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GIDIFASKNFH 11

DB 171 GMSVPAKNFH 181

RESULT 9

T39029

hypothetical protein SPAC6C3.05 - fission yeast (*Schizosaccharomyces pombe*)

C/Species: *Schizosaccharomyces pombe*

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C/Accession: T39029

R/Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, February 1996

A/Reference number: Z21750

A/Accession: T39029

A/Status: Preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-269 <DEV>

A/Cross-references: EMBL:Z69731; PIDN:CA840280.1; GSPDB:GN00066; SPDB:SPAC6C3.05

A/Experimental source: strain 972h; cosmid c6C3

C/Genetics:

A/Gene: SPDB:SPAC6C3.05

A/Map position: 1

Query Match

Best Local Similarity 51.9%; Score 41; DB 2; Length 269;

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHL 14

DB 202 GLEISASKNFLEK 215

RESULT 10

S55645

transcription control protein 50 - equine herpesvirus 2

C/Species: equine herpesvirus 2

C/Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 08-Oct-1999

C/Accession: S55645

R/Reid, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.

J. Mol. Biol. 249, 520-528, 1995

A/Title: The DNA sequence of equine herpesvirus 2.

A/Reference number: S55594; MUID:95302501; PMID:7783207

A/Accession: S55645

A/Status: Preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-630 <TEL>

A/Cross-references: GB:U20824; NID:9695172; PIDN:AA013038.1; PID:9695223

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

Query Match

Best Local Similarity 51.9%; Score 41; DB 2; Length 630;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 IDIFASKNFHL 15

DB 614 LDIFSLNHLHRS 627

RESULT 11

B90589

sugar ABC transporter permease protein [imported] - *Mycoplasma pulmonis* (strain UAB C

C/Species: *Mycoplasma pulmonis*

C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

C/Accession: B90589

R/Chambaud, I.; Hellig, R.; Ferris, S.; Barde, V.; Samson, D.; Gallison, F.; Moszer, I

Nucleic Acids Res. 29, 2145-2153, 2001

A/Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma p*

A/Reference number: A99512; MUID:21267105; PMID:11353084

A/Accession: B90589

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-598 <NUR>

A/Cross-references: GB:AL445566; PID:G1409003; PIDN:CA031791.1; GSPDB:GN00153

A/Experimental source: strain UAB CTIP

C/Genetics:

A/Gene: MYPV 6180

A/Genetic code: SGC3

Query Match

Best Local Similarity 50.6%; Score 40; DB 2; Length 598;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GIDIFASKNFH 11

DB 383 GINIFSKNFH 393

RESULT 12

S64506

protein kinase BUB1 (EC 2.7.1.-), checkpoint-associated - yeast (*Saccharomyces cerevis*

N/Alternate names: protein G7542; protein YGR188C

C/Species: *Saccharomyces cerevisiae*

C/Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 21-Jul-2000

C/Accession: S64506; A56354; S50224

R/Arroyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M.I.; Sanchez-Perez, M.; Nombela, C.

submitted to the Protein Sequence Database, May 1996

A/Reference number: S64499

A/Accession: S64506

A/Molecule type: DNA

A/Residues: 1-1021 <ABR>

A/Experimental source: strain S288C

R/Roberts, B.T.; Farr, K.A.; Hoyt, M.A.

Mol. Cell. Biol. 14, 8282-8291, 1994

A/Title: The *Saccharomyces cerevisiae* checkpoint gene BUB1 encodes a novel protein kin

A/Reference number: A56354; MUID:95059057; PMID:7969164

A/Accession: A56354

A/Status: Preliminary

A:Molecule type: DNA
 A:Residues: 1-530, 'V', 532-1021 <ROB>
 A:Cross-references: GB:I32027; NID:G475127; PIDN:AAA64894.1; PID:G475128
 C:Genetics:
 A:Gene: SGD:BUB1
 A:Cross-references: SGD:S0003420; MIPS:YGR188c
 A:Map position: 7R
 C:Keywords: autophosphorylation; cell division control; phosphoprotein; phosphotransfera

Query Match
 Best Local Similarity 50.6%; Score 40; DB 2; Length 1021;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 DIFASKNFHLQ 15
 DB 115 INLFNNFHESEN 128

RESULT 13
 T43075
 Hypothetical protein - Lactococcus lactis plasmid pMRC01

C:Species: Lactococcus lactis
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T43075
 R:Dougherty, B.A.; Hill, C.; Weidman, J.F.; Richardson, D.R.; Venter, J.C.; Ross, R.P.
 Mol. Microbiol. 29, 1029-1038, 1998

A:Title: Sequence and analysis of the 60 kb conjugative, bacteriocin-producing plasmid F

A:Reference number: Z22314
 A:Accession: T43075
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-64 <DOU>

A:Cross-references: EMBL:AE001272; PIDN:AACS6026.1

A:Experimental source: strain DPC3147

C:Genetics:
 A:Genome: plasmid pMRC01
 A:Note: ORF0002

Query Match
 Best Local Similarity 49.4%; Score 39; DB 2; Length 64;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHLQ 13
 DB 44 GILIFLKNFHVQ 56

RESULT 14

S73916
 MG103 homolog K04_crf280 - Mycoplasma pneumoniae (strain ATCC 29342)

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999

C:Accession: S73916
 R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A:Reference number: S73327; MUID:97105885; PMID:8948633

A:Accession: S73916
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-280 <HIM>
 A:Cross-references: EMBL:AE000058; GB:U00089; NID:G1674291; PIDN:AA96238.1; PID:G167429

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:Genetics:
 A:Genetic code: SGC3

Query Match
 Best Local Similarity 49.4%; Score 39; DB 2; Length 280;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

DB 114 DILHSSNFHLQ 124

RESULT 15

D64211
 Hypothetical protein MG103 - Mycoplasma genitalium

C:Species: Mycoplasma genitalium

C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 07-Dec-1999

C:Accession: D64211
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.

M.; Fuhmann, J.; Nguyen, D.; Usterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,

, C.A.; Venter, J.C.
 Science 270, 397-403, 1995

A:Title: The minimal gene complement of Mycoplasma genitalium.

A:Reference number: A64200; MUID:96026346; PMID:7569993

A:Accession: D64211
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-280 <TIGR>

A:Cross-references: GB:U39689; GB:I43967; NID:G1045773; PID:G1045781; TIGR:MG103

A:Experimental source: strain G-37

C:Genetics:
 A:Genetic code: SGC3

Query Match
 Best Local Similarity 49.4%; Score 39; DB 2; Length 280;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 DIFASKNFHLQ 13
 DB 114 DLINSNFHLQ 124

Search completed: April 20, 2003, 13:15:54
 Job time: 8.07895 secs

GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524d-120

Perfect score: 79

Sequence: 1 GIDIFASKNFHLQKN 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	514	1	MPA2_CRYVA
2	43	54.4	419	1	YIY9_SCHPO
3	42	53.2	357	1	VG58_HSVSA
4	41	51.9	269	1	YD55_SCHPO
5	40	50.6	1021	1	BUB1_YEAST
6	39	49.4	273	1	BCHL_CHLAU
7	39	49.4	280	1	Y103_MYCCE
8	39	49.4	280	1	Y103_MYCCE
9	39	49.4	349	1	Y44F_SCHPO
10	39	49.4	437	1	PERG_LACDL
11	38	48.1	509	1	DLDH_PIG
12	38	48.1	824	1	CC27_HUMAN
13	37	46.8	155	1	IL12_RAT
14	37	46.8	223	1	PURQ_LACIA
15	37	46.8	352	1	SMP_SERMA
16	37	46.8	509	1	DLDH_HUMAN
17	37	46.8	509	1	DLDH_HUMAN
18	37	46.8	701	1	Y363_BUCAI
19	37	46.8	1031	1	YD69_SCHPO
20	36	45.6	137	1	MESC_LETME
21	36	45.6	153	1	IL2_HUMAN
22	36	45.6	154	1	IL2_CERO
23	36	45.6	154	1	IL2_MACNU
24	36	45.6	208	1	PSB2_DROME
25	36	45.6	234	1	A29B_DROME
26	36	45.6	284	1	YMX7_YEAST
27	36	45.6	330	1	UL16_HSVSA
28	36	45.6	331	1	GPDA_CLOPE
29	36	45.6	358	1	Y39L_YEAST
30	36	45.6	412	1	FTS2_ENTFA
31	36	45.6	620	1	YG18_YEAST
32	36	45.6	672	1	MIS6_SCHPO
33	36	45.6	706	1	HDAL_YEAST

34	36	45.6	752	1	CO2_HUMAN	P06681	homo sapien
35	36	45.6	768	1	YB23_HUMAN	O99117	homo sapien
36	36	45.6	768	1	EXG1_COCCA	P49426	cochliobol
37	35.5	44.9	342	1	QVEA_STREN	O97915	streptococc
38	35.5	44.9	342	1	QVEA_STREN	O99249	streptococc
39	35	44.3	132	1	TVAA_MOUSE	P06323	mus musculu
40	35	44.3	140	1	NDK_RHOSU	P56553	rhodovulum
41	35	44.3	147	1	YBAK_RACSU	P50862	bacillus su
42	35	44.3	155	1	VG06_YACCC	P21027	vaccinia vi
43	35	44.3	165	1	VG06_YARV	P32996	variola vir
44	35	44.3	170	1	YHE3_YEAST	P38728	saccharomyc
45	35	44.3	194	1	YCEF_ECO57	P58626	escherichia

ALIGNMENTS

RESULT 1
MPA2_CRYVA STANDARD; PRT; 514 AA.
AC P43212;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)
DE (Major pollen allergen Cry j 2) (Cry j II).
OS Cryptomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OX NCBI_TaxID=3369;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=95010777; PubMed=7926035;
RA Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,
RA Usui M., Kurimoto M.,
RT "Molecular cloning of the second major allergen, Cry j II, from
RT Japanese cedar pollen."
RL FEBS Lett. 353:124-128(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=94271186; PubMed=8002972;
RA Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.,
RT "cDNA cloning and expression of Cry j II the second major allergen of
RL Japanese cedar pollen."
RL Biochem. Biophys. Res. Commun. 201:1021-1028(1994).
RN [3]
RP SEQUENCE OF 55-64.
RX MEDLINE=90342988; PubMed=2382797;
RA Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matubasi T.,
RT "Identification of the second major allergen of Japanese cedar
RT pollen."
RL Allergy 45:309-312(1990).
CC -I- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -I- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).
CC -I- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
(POLYGALACTURONASES).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; D37765; BAA07021.1; -;
CC EMBL; D29772; BAA06172.1; -;
CC HSSP; P26509; 1BHE.
DR InterPro: IPR000743; GH28.
DR Pfam: PF00295; Glyco_hydro_28; 1.

DR PROSITE; PS00502; POLYGALACTURONASE, 1.
 KM Hydroxylase; Glycosidase; Cell wall; signal; Zymogen; Fruit ripening;
 KW Amyloplast; Glycoprotein; Allergen.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPER 23 45 POTENTIAL.
 FT CHAIN 46 433 POSSIBLE POLYGALACTURONASE.
 FT ACT SITE 434 514 POTENTIAL.
 FT CARBOHYD 278 278 PROBABLE.
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 472 472 F -> L (IN REF. 2).
 FT CONFLICT 5 5 V -> L (IN REF. 2).
 FT CONFLICT 12 12 Q -> VV (IN REF. 2).
 FT CONFLICT 34 35 Q -> K (IN REF. 2).
 FT CONFLICT 37 37 K -> N (IN REF. 2).
 FT CONFLICT 88 88 N -> S (IN REF. 2).
 FT CONFLICT 98 98 K -> E (IN REF. 2).
 FT CONFLICT 451 451 G -> R (IN REF. 2).
 FT CONFLICT 454 454 M -> I (IN REF. 2).
 FT CONFLICT 504 504 R -> C (IN REF. 2).
 FT CONFLICT 507 507
 SQ SEQUENCE 514 AA; 56645 MW; 624611C3FA8D6302 CRC64;

Query Match 100.0%; Score 79; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHLQKN 15
 DB 235 GIDIFASKNFHLQKN 249

RESULT 2
 YTY9_SCHPO STANDARD; PRT; 419 AA.
 AC Q9YU14;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C22F8.09 in chromosome I.
 GN SPAC22F8.09.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 CC NCBI_Taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=21848401; PubMed=11859360;
 RC STRAIN=972;
 RX Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 Mooney P., Moule S., Mungall K., Murphy L., Miblett D., Odell C.,
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 Woodward J., Volkart G., Aert R., Robben J., Grymptre B.,
 Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs W., Filtz C., Holzer E., Moestl D., Hilbert H.,
 Borzym K., Langer I., Beck A., Lehach H., Reinhardt R., Pohl T.M.,
 Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
 Goffeau A., Cadieu E., Driano S., Gloux S., Lelaure V., Mottier S.,
 Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huest S.M.,
 Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 Dominguez A., Revuelta J.U., Moreno S., Armstrong J., Forsburg S.L.,
 Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO THE GLTSCR2 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AL109831; CAB52719.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 419 AA; 48129 MW; 9C2A72B639E61DAA CRC64;

QY 5 FASKNFHLQKN 15
 DB 407 YSHKDFHLQKN 417

RESULT 3
 VG58_HSVSA STANDARD; PRT; 357 AA.
 AC D01053;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Gene 58 protein.
 GN 58 OR EBFP5.
 OS Herpesvirus saimiri (strain 11).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 CC NCBI_Taxid=10383;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=92333688; PubMed=1321287;
 RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
 Newman C., Wiltmann S., Craxton M.A., Coleman H., Fleckenstein B.,
 Honess R.W.;
 RT "Primary structure of the herpesvirus saimiri genome."
 RT J. Virol. 66:5047-5058(1992).
 RL [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92230228; PubMed=1314457;
 RA Nicholas J., Cameron K.R., Coleman H., Newman C., Honess R.W.;
 RT "Analysis of nucleotide sequence of the rightmost 43 kbp of
 RT herpesvirus saimiri (HVS) L-DNA: general conservation of genetic
 RT organization between HVS and Epstein-Barr virus."
 RL Virology 188:296-310(1992).
 CC -1- SIMILARITY: TO EBV BMRF2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X64346; CAA45681.1; -
 DR EMBL; M86409; AAA46134.1; -
 DR PIR; B36812; Q0BEP3.
 SQ SEQUENCE 357 AA; 40526 MW; 390F01A29F9F56B1 CRC64;

Query Match 53.2%; Score 42; DB 1; Length 357;
 Best Local Similarity 63.6%; Pred. No. 4;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GIDIFASKNFH 11
Db 171 GNGVFAKNFH 181

RESULT 4

YD55_SCHPO STANDARD; PRT; 269 AA.
AC Q10308;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C6C3.05 in chromosome I.
GN SPAC6C3.05.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gantley S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean U.,
RA Mooney P., Moulis S., Mungall K., Murphy L., Nidlett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Sefton J., Simmonds M., Squares R., Stevens K., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicakeet G., Aert R., Robben J., Grymoprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z69731; CAB40280.1; -
DR Hypothetical protein.
KW SEQUENCE 269 AA; 31234 MW; 0C3DF87A138CA5BE CRC64;

Query Match 51.9%; Score 41; DB 1; Length 269;
Best Local Similarity 57.1%; Pred. NO. 4.4;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 GIDIFASKNFHLOK 14
Db 202 GUEISASKNFLEK 215

RESULT 5
BUB1_YEAST

ID BUB1_YEAST STANDARD; PRT; 1021 AA.
AC P41695;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Checkpoint serine/threonine-protein kinase BUB1 (EC 2.7.1.-).
GN BUB1 OR YGR188C OR G7542.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97279231; PubMed=7969164;
RA Roberts B.T., Farr K.A., Hoyt M.A.;
RT "The Saccharomyces cerevisiae checkpoint gene BUB1 encodes a novel
RL Mol. Cell. Biol. 14:8282-8291 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97279231; PubMed=9133739;
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
RA Nombela C.;
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
RL Yeast 13:357-363 (1997).
CC -1- FUNCTION: INVOLVED IN CELL CYCLE CHECKPOINT ENFORCEMENT. CATALYZES
CC THE PHOSPHORYLATION OF BUB3 AND ITS AUTOPOHOSPHORYLATION.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: AUTOPOHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC BUB1 SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 1 CD1 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L32027; AAA64894.1; -
DR EMBL; 272973; CAA97214.1; -
DR EMBL; X99074; CAA67524.1; -
DR SGD; S0003420; BUB1.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR Prodom: PPO00001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Phosphorylation; Nuclear protein.
FT DOMAIN 55 304
FT NP_BIND 705 1021 PROTEIN KINASE.
FT BINDING 711 719 ATP (BY SIMILARITY).
FT ACT_SITE 833 833 ATP.
FT MUTAGEN 733 733 BY SIMILARITY.
FT CONFLICT 531 531 D -> V (IN REF. 1).
SQ SEQUENCE 1021 AA; 117868 MW; 6D76FC980775D3F9 CRC64;

Query Match 50.6%; Score 40; DB 1; Length 1021;
Best Local Similarity 50.0%; Pred. NO. 28;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 IDIFASKNFHLOK 15
Db 115 INLFASKNFHSEN 128

RESULT 6

BCHL_CHLNU STANDARD; PRT: 273 AA.

ID BCBL_CHLNU

AC Q9F6X3; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Light-independent protochlorophyllide reductase iron-sulfur ATP-binding protein (EC 1.18.-.-) (LIPOR subunit L) (DOR subunit L).

GN BCLH.

OS Chloroflexus aurantiacus.

OC Bacteria; Chloroflexi; Chloroflexales; Chloroflexaceae; Chloroflexus.

OX NCBI_TaxID=1108;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20433268; PubMed=10976061;

RA Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.E.;

RT "Molecular evidence for the early evolution of photosynthesis.";

RL Science 289:1724-1730(2000).

CC -1- FUNCTION: Uses Mg-ATP and reduced ferredoxin to reduce ring D of protochlorophyllide (Pchl_{id}) to form chlorophyllide a (Chl_{id}) (By similarity). This reaction is light-independent.

CC -1- PATHWAY: Light-independent bacteriochlorophyll biosynthesis.

CC -1- SUBUNIT: Protochlorophyllide reductase is thought to be composed of three subunits: bchl, bchm and bchn. Homodimer of bchl subunit (By similarity).

CC -1- SIMILARITY: BELONGS TO THE NIFH / BCLH / CHL FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AF288460; MAG5121.1; -

DR HSSP: P00459; 1FP6.

DR InterPro: IPR000392; NitrogenaseI.

DR Pfam: PR00142; fer4_nifh; 1.

DR PRINTS: PR00091; NITROGNASEI.

DR TIGRPFAMs: TIGR01281; DPOR_DCHL; 1.

DR PROSITE: PS00746; NIFH_FRXC_1; 1.

DR PROSITE: PS00692; NIFH_FRXC_2; 1.

KW Oxidoreductase; Photosynthesis; Bacteriochlorophyll biosynthesis;

KW ATP-binding; Iron-sulfur; 4Fe-4S.

FT NP BIND 9 16 ATP (POTENTIAL).

FT METAL 97 97 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

FT METAL 131 131 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

SQ SEQUENCE 273 AA; 29439 MW; 4F7A9A97C5343B5B CRC64;

Query Match 49.4%; Score 39; DB 1; Length 273;

Best Local Similarity 58.3%; Pred. No. 10;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 IDIFASKNFHQ 13

DB 57 IDVLDSVNFHLE 68

RESULT 7

ID Y103 MYCPE STANDARD; PRT: 280 AA.

AC P47349; Q49293; 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein MG103.

GN MG103.

OS Mycoplasma genitalium.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2097;

RESULT 8

ID Y103 MYCPE STANDARD; PRT: 280 AA.

AC P75530; 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein MG103 homolog (K04_orf280).

GN MEN241 OR MP590.

OS Mycoplasma pneumoniae.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2104;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=ATCC 29342 / M129;

RA MEDLINE=97105885; PubMed=8948633;

RA Heremans R., Hilbert H., Pliagens H., Pirk E., Li B.-C.,

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";

RL Nucleic Acids Res. 24:4420-4449(1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC EMBL: U39690; AAC71321.1; -

DR EMBL: U02170; AAD12452.1; -

DR TIGR: MG103; -

DR InterPro: IPR005802; DUF199.

DR Pfam: PF02650; DUF199; 1.

DR TIGRPFAMs: TIGR00647; MG103; 1.

KW Hypothetical protein; Complete proteome.

FT CONFLICT 86 86 L -> F (IN REF. 2).

SQ SEQUENCE 280 AA; 32864 MW; 022BC4D71D75A2D CRC64;

Query Match 49.4%; Score 39; DB 1; Length 280;

Best Local Similarity 63.6%; Pred. No. 11;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 DIFASKNFHQ 13

DB 114 DLINNNFHQ 124

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AE000058; AAB96238.1; -
 DR InterPro; IPR003802; DUF199.
 DR Pfam; PF02650; DUF199; 1.
 DR TIGRPFAMs; TIGR00647; MG103; 1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 280 AA; 32194 MW; 5842D83CBFC26AAA CRC64;

Query Match 49.4%; Score 39; DB 1; Length 280;
 Best Local Similarity 63.6%; Pred. No. 11;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 DIFASKNFHLQ 13
 Db 114 DLHSSNPHLQ 124

RESULT 9
 YAAF SCHPO STANDARD; PRT; 349 AA.
 AC Q09732;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C31A2.15c in chromosome 1.
 GN SPAC31A2.15c.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 RN NCB1_TaxID=4896;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Stimmings M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitlehead S.,
 RA Woodward J., Volkert G., Aert R., Robben J., Grymporter B.,
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Punnett B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mortier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.U., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrall B.G., Nure P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; Z50113; CA90473.1; -
 DR Hypothetical protein.
 SQ SEQUENCE 349 AA; 40938 MW; 3143B4EB763D280A CRC64;

Query Match 49.4%; Score 39; DB 1; Length 349;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 IDIFASKNFHLQXN 15
 Db 143 VDSFLOKNFPMVKN 156

RESULT 10
 PEPG LACDL STANDARD; PRT; 437 AA.
 AC P94869;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aminopeptidase G (EC 3.4.22.-).
 GN PEPG.
 OS Lactobacillus delbrueckii (subsp. lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus.
 RN NCB1_TaxID=29397;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 7290;
 RX MEDLINE=97195798; PubMed=9043129;
 RA Klein J.R., Schick J., Henrich B., Plapp R.,
 RA "Lactobacillus delbrueckii subsp. lactis DSM7290 pepg gene encodes a
 RT novel cysteine aminopeptidase.";
 RL Microbiology 143:527-537(1997).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; Z71782; CA96465.1; -
 DR HSSP; Q13667; 2CB5.
 DR MEROPS; C01.089; -.
 DR InterPro; IPR004134; Pept C1-like.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHprol-acetile.
 DR Pfam; PF03051; Pept C1-like; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PROSITE; PS00339; THOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THOL_PROTEASE_HIS; 1.
 DR PROSITE; PS00640; THOL_PROTEASE_ASN; FALSE-NEG.
 KM Hydroxylase; Thiol protease; Aminopeptidase.
 FT ACT SITE 70
 FT ACT SITE 361
 FT ACT SITE 382
 FT ACT SITE 382
 SQ SEQUENCE 437 AA; 49705 MW; 2099C62E57181FC5 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 437;
 Best Local Similarity 75.0%; Pred. No. 17;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 KNFHLQXN 15
 Db 234 KNFHLQXN 241

RESULT 11

ID	DLID	PID	STANDARD	PRT	509 AA
AC	P09623				
DT	01-MAR-1989	(Rel. 10, Created)			
DT	01-MAR-1989	(Rel. 10, Last sequence update)			
DT	15-JUL-1999	(Rel. 38, Last annotation update)			
DE	15-JUL-1999	(Rel. 38, Last annotation update)			
GN	DLID OR LAD				
OS	Sus scrofa (Pig)				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus				
OX	NCBI_Taxid=9823				
RN	SEQUENCE FROM N.A.				
RP	MEDLINE=88087005; PubMed=3693355;				
RA	Otlakowski G., Robinson B.H.;				
RT	"Isolation and sequence determination of cDNA clones for porcine and human liponamide dehydrogenase. Homology to other disulfide oxidoreductases."				
RL	J. Biol. Chem. 262:17313-17318(1987).				
CC	-1- FUNCTION: LIPONAMIDE DEHYDROGENASE IS A COMPONENT OF THE GLYCINE CLEAVAGE SYSTEM AS WELL AS OF THE ALPHA-KETOACID DEHYDROGENASE COMPLEXES.				
CC	-1- CATALYTIC ACTIVITY: Dihydroliponamide + NAD(+) = Liponamide + NADH.				
CC	-1- COFACTOR: FAD.				
CC	-1- SUBUNIT: HOMODIMER.				
CC	-1- SUBCELLULAR LOCATION: Mitochondrial matrix.				
CC	-1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.				
CC	-1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE OXIDOREDUCTASES CLASS-1.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; J03489; AAA31069.1; -				
DR	PIR; A28448; DEPGP.				
DR	HSSP; P18925; 3LAD.				
DR	InterPro; IPR001337; FAD pyr redox.				
DR	InterPro; IPR001100; Pyr_redox.				
DR	InterPro; IPR004099; Pyr_redox_dim.				
DR	Pfam; PF00070; Pyr_redox_1.				
DR	Pfam; PF02852; Pyr_redox_dim; 1.				
DR	PRINTS; PR00368; FADPFR.				
DR	PRINTS; PR00411; PNDRTASE1.				
DR	ProDom; PD000139; FAD pyr redox; 1.				
DR	PROSITE; PS00076; PYRIDINE-REDOX_1; 1.				
KW	Redox-active center; Oxidoreductase; NAD; Flavoprotein; FAD; Mitochondrion; Transil peptide.				
KW	TRANSIT				
FT	CHAIN	1	35		
FT	NP_BIND	36	509		
FT	NP_BIND	43	73		
FT	DISULFID	80	85		
FT	NP_BIND	345	355		
FT	ACT_SITE	487	487		
FT	SEQUENCE	509 AA; 54185 MW; 38A0469FED071300 CRC64;			

Query Match 48.1%; Score 38; DB 1; Length 509; Best Local Similarity 62.5%; Pred. No. 30; Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

CC27 HUMAN

ID	CC27 HUMAN	STANDARD	PRT	824 AA
AC	P30250; O16349; Q96F35;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Protein CDC27H (Cell division cycle protein 27 homolog) (H-NUC).			
GN	CDC27.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=94052097; PubMed=8234252;			
RA	Tyngendreich S., Boguet M.S., Seldin M., Hietel P.A.;			
RT	"Linking yeast genetics to mammalian genomes: identification and mapping of the human homolog of CDC27 via the expressed sequence tag (EST) data base."			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:10031-10035(1993).			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=95275739; PubMed=7756179;			
RA	Chen P.L., Ueng Y.C., Duffee T., Chen K.C., Yang-Feng T., Lee W.H.;			
RT	"Identification of a human homologue of yeast nuc2 which interacts with the retinoblastoma protein in a specific manner."			
RL	Cell Growth Differ. 6:199-210(1995).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISUB=Uterus;			
RA	Strausberg R.;			
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.			
CC	-1- SUBUNIT: INTERACTS WITH RB.			
CC	-1- SUBCELLULAR LOCATION: Nucleus.			
CC	-1- SIMILARITY: BELONGS TO THE CDC27/NUC2 FAMILY.			
CC	-1- SIMILARITY: CONTAINS 9 TPR REPEATS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U00001; AAA60471.1; -			
DR	EMBL; S78234; AAB34378.1; -			
DR	EMBL; BC011656; AAH11656.1; -			
DR	Genew; HGNC:1728; CDC27.			
DR	MTM; 116946; -			
DR	InterPro; IPR001440; TPR.			
DR	Pfam; PF00515; TPR; 8.			
DR	SMART; SMO0028; TPR; 7.			
KW	Repeat; TPR repeat; Nuclear protein.			
FT	REPEAT	84	114	
FT	REPEAT	115	148	
FT	REPEAT	499	532	
FT	REPEAT	567	600	
FT	REPEAT	602	634	
FT	REPEAT	635	668	
FT	REPEAT	670	702	
FT	REPEAT	704	736	
FT	REPEAT	737	770	
FT	REPEAT	319	319	
FT	CONFLICT	403	403	
FT	CONFLICT	460	460	
FT	CONFLICT	715	715	
FT	SEQUENCE	824 AA; 91867 MW; E6C8F59C1EFDCBA CRC64;		

Query Match 48.1%; Score 38; DB 1; Length 824; Best Local Similarity 40.0%; Pred. No. 51; Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 1 GIDIFASKNFHLQXN 15
 DB 535 GMEIYSTLMLHOKD 549

RESULT 13

IL2_RAT
 ID IL2_RAT STANDARD; PRT; 155 AA.
 AC P17108;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Interleukin-2 precursor (IL-2) (T-cell growth factor) (TCGF).
 GN IL2 OR IL-2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_Taxid=10116;
 [1]

SEQUENCE FROM N.A.
 RX MEDLINE=89339608; PubMed=2788130;
 RA McKnight A.J., Mason D.W., Barclay A.N.;
 RT "Sequence of rat interleukin 2 and anomalous binding of a mouse
 RT interleukin 2 cDNA probe to rat MHC class II-associated invariant
 RT chain mRNA."
 RL Immunogenetics 30:145-147(1989).
 CC -1- FUNCTION: PRODUCED BY T-CELLS IN RESPONSE TO ANTIGENIC OR
 CC MYTOGENIC STIMULATION, THIS PROTEIN IS REQUIRED FOR T-CELL
 CC PROLIFERATION AND OTHER ACTIVITIES CRUCIAL TO REGULATION OF THE
 CC IMMUNE RESPONSE. CAN STIMULATE B CELLS, MONOCYTES, LYMPHOKINE-
 CC ACTIVATED KILLER CELLS, NATURAL KILLER CELLS, AND GLIOMA CELLS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE IL-2 FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M22899; AAA41427.1; -
 CC DR PIR; A31278; A31278.
 CC DR PIR; A45882; A45882.
 CC DR HSSP; P01585; SINK.
 CC DR InterPro; IPR000779; Interleukin-2.
 CC DR Pfam; PF00715; IL2; 1.
 CC DR PRINTS; PR00265; INTERLEUKIN2.
 CC DR ProDom; PD003649; Interleukin-2; 1.
 CC DR SMART; SM00189; IL2; 1.
 CC DR PROSITE; PS00424; INTERLEUKIN_2; 1.
 CC KW Cytokine; Glycoprotein; Immune response; Signal; Growth factor;
 CC T-cell.
 CC FT SIGNAL 1 20 BY SIMILARITY.
 CC FT CHAIN 21 155 INTERLEUKIN-2.
 CC FT CARBOHYD 23 23 O-LINKED (GALNAc...) (BY SIMILARITY).
 CC FT DISULFID 78 126 BY SIMILARITY.
 CC SQ SEQUENCE 155 AA; 17632 MW; 67A8554A73BF30A0 CRC64;

Query Match 46.8%; Score 37; DB 1; Length 155;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 IDIFASKNFHLQ 13
 DB 90 LDLTQSKSFHLE 101

RESULT 14
 ID PURO_LACLA STANDARD; PRT; 223 AA.
 AC O9CFE7;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphoribosylformylglycinamide synthase I (EC 6.3.5.3) (PGAM
 DE synthase I).
 GN PURO OR IL1531.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 NCBI_Taxid=1360;
 [1]
 SEQUENCE FROM N.A.
 RX STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jallion O., Malarne K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403."
 RL Genome Res. 11:731-753(2001).
 CC -1- CATALYTIC ACTIVITY: ATP + N(2)-formyl-N(1)-(5-phospho-D-
 CC ribosyl)glycinamide + L-glutamine + H(2)O = ADP + phosphate + 2-
 CC (formamido)-N(1)-(5-phospho-D-ribosyl)acetamide + L-glutamate.
 CC -1- PATHWAY: De novo purine biosynthesis; fourth step.
 CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, PURO AND PURL.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; AE006384; AK05629.1; -
 DR InterPro; IPR000991; GATASE_1.
 DR InterPro; IPR002818; ThJd.
 DR Pfam; PF01965; ThJd; 1.
 DR PROSITE; PS00442; GATASE_TYPE_I; FALSE NEG.
 KW Purine biosynthesis; Ligase; Glutamine amidotransferase;
 KW Complete proteome.
 FT ACT_SITE 86 GATASE (BY SIMILARITY).
 FT SEQUENCE 223 AA; 24205 MW; A8FED3AE08A1262 CRC64;

Query Match 46.8%; Score 37; DB 1; Length 223;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 8; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

OY 1 GIDIFAS-KNF 10
 DB 209 GVDLFASVKNF 220

RESULT 15
 SMP_SERMA
 ID SMP_SERMA STANDARD; PRT; 352 AA.
 AC Q06517;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Extracellular minor metalloprotease precursor (EC 3.4.24.-).
 GN SMP.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 NCBI_Taxid=615;
 [1]

SEQUENCE FROM N.A., AND SEQUENCE OF 53-57.
 RC STRAIN=ATCC 21074;
 RX MEDLINE=93194081; PubMed=8449415;
 RA Kwon Y.T., Lee H.H., Rho H.N.;
 RT "Cloning, sequencing, and expression of a minor protease-encoding
 RT gene from Serratia marcescens ATCC21074.";

```

RL Gene 125:75-80 (1993).
CC -1- COFACTOR: BINDS 1 ZINC ION.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THE OPTIMAL TEMPERATURE AND PH FOR ENZYME ACTIVITY
CC ARE 50 DEGREES CELSIUS AND PH 8.0.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59854; AAA26553.1; -.
CC PIR: J0153; J0153.
CC HSSP: P05806; INPC.
DR MEROPS; M04.UFW; -.
DR InterPro; IPR001570; Peptidase_M4.
DR InterPro; IPR000130; Zn_MTPeptidase.
DR Pfam; PF02868; Peptidase_M4_C; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KM Hydrolase; Metalloprotease; zinc; Signal; Zymogen.
FT SIGNAL 1
FT PROBE 52
FT CHAIN 53 352 ACTIVATION PEPTIDE.
FT METAL 160 160 EXTRACELLULAR MINOR METALLOPROTEASE.
FT ACT_SITE 161 161 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 164 164 BY SIMILARITY.
FT METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 184 184 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 262 262 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 352 AA; 38515 MW; D10IC956426B163C CRC64;

```

```

Query Match 46.8%; Score 37; DB 1; Length 352;
Best Local Similarity 57.1%; Pred. No. 31;
Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

```

```

QY 3 DIFAS--KNFHLOK 14
DB 188 DVFGSLVKQFHLOQ 201

```

```

Search completed: April 20, 2003, 13:07:43
Job time : 4.92105 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 12.6711 Seconds
(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524D-120
Perfect score: 79
Sequence: 1 GIDIFASKNFHLQKN 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues
Total number of hits satisfying chosen parameters:

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing First 45 summaries

```

1:  sp_archaea:*
2:  sp_bacteria:*
3:  sp_fungi:*
4:  sp_human:*
5:  sp_invertebrate:*
6:  sp_mammal:*
7:  sp_mhc:*
8:  sp_organelle:*
9:  sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriopl:*
17: sp_archaeopl:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Match	Score	Query	length	DB	ID	Description
1		54	68.4	507	10	Q9FY19	Q9FY19 juniperus a
2		45	57.0	503	13	Q9W6Y2	Q9W6Y2 oncorhynch
3		43	54.4	252	2	Q9AHX1	Q9AHX1 carsonella
4		43	54.4	1360	5	O15807	O15807 cryptospori
5		42	53.4	1431	5	Q9XYH6	Q9XYH6 cryptospori
6		42	53.2	156	16	Q9POJ1	Q9POJ1 ureaplasma
7		42	53.2	305	5	O45343	O45343 caenorhabd
8		41	51.9	630	12	O666E2	O666E2 equine hepr
9		40	50.6	251	10	O81407	O81407 sinapis arv
10		40	50.6	445	16	Q8XJS9	Q8XJS9 clostridium
11		40	50.6	598	16	Q98PV1	Q98PV1 mycoplasma
12		40	50.6	859	5	Q9VRL4	Q9VRL4 drosophila
13		39	50.0	490	10	Q9XEY7	Q9XEY7 oryza sativ
14		39	49.4	64	2	O87205	O87205 lactococcus
15		39	49.4	122	12	O11321	O11321 molluscum c
16		39	49.4	204	16	Q8FSN9	Q8FSN9 fusobacteri

17	39	49.4	335	4	Q9Y382	homo sapien
18	39	49.4	353	16	Q9WY02	Q9WY02 thermotoga
19	39	49.4	375	10	Q9SSR8	Q9SSR8 arabidopsis
20	39	49.4	401	10	Q24554	Q24554 zinnia elec
21	39	49.4	430	10	Q9AW73	Q9AW73 giillardia
22	39	49.4	565	12	Q98205	Q98205 molluscum c
23	39	49.4	623	16	Q8R669	Q8R669 fusobacteri
24	39	49.4	904	5	Q9U011	Q9U011 plasmodium
25	39	49.4	1073	5	Q9GDN3	Q9GDN3 neurospora
26	38	48.1	34	5	Q8TKN5	Q8TKN5 caenorhabdi
27	38	48.1	161	3	Q9P3N1	Q9P3N1 neurospora
28	38	48.1	170	2	Q32478	Q32478 rhodospira
29	38	48.1	248	2	Q93UB6	Q93UB6 carsonella
30	38	48.1	248	2	Q93UB8	Q93UB8 carsonella
31	38	48.1	253	2	Q93UB6	Q93UB6 carsonella
32	38	48.1	258	5	Q9VR7	Q9VR7 dirosophila
33	38	48.1	324	4	Q9Y310	Q9Y310 homo sapien
34	38	48.1	348	12	Q57103	Q57103 monkeypox v
35	38	48.1	348	12	Q57108	Q57108 monkeypox v
36	38	48.1	348	12	Q57277	Q57277 monkeypox v
37	38	48.1	349	12	Q57099	Q57099 monkeypox v
38	38	48.1	349	12	Q57100	Q57100 monkeypox v
39	38	48.1	349	12	Q57101	Q57101 monkeypox v
40	38	48.1	349	12	Q57102	Q57102 monkeypox v
41	38	48.1	349	12	Q57291	Q57291 monkeypox v
42	38	48.1	435	16	Q9Z8Y7	Q9Z8Y7 chlamydia p
43	38	48.1	506	5	Q9V5X2	Q9V5X2 drosohila
44	38	48.1	615	16	Q8YKX5	Q8YKX5 anabaena sp
45	38	48.1	625	16	Q9YJ52	Q9YJ52 clostridium

ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	507 AA.
09FY19			
ID 09FY19			
AC 09FY19;			
DT 01-MAR-2001 (TREMBLrel. 16, Created)			
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE Pollen major allergen 2 protein precursor.			
UNA2.			
OS Juniperus ashei (Ozark white cedar).			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.			
OX NCBI_TaxId=13101;			
RP [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=MALE POLLEN;			
RX MEDLINE=20403896; PubMed=10944464;			
RA Yokoyama M., Miyahara M., Shimizu K., Kino K., Tsunoo H.;			
RT "Purification, Identification and cDNA cloning of Jun a 2, the second			
RL major allergen of mountain cedar pollen.";			
RL Biochem. Biophys. Res. Commun. 275:195-202 (2000).			
-1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES			
CC (POLYGALACTURONASES).			
CC EMBL: AJ1404653; CAC05582.1; ..			
DR HSSP; P26509.1BHE.			
DR InterPro; IPR000743; GH28.			
DR InterPro; IPR000408; Reg_chir_condens.			
DR Pfam; PF00295; Glyco_hydro_28; 1.			
DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.			
DR PROSITE; PS00626; RCCL_2; UNKNOWN_1.			
KW Cell wall; Glycosidase; Hydrolase; Signal.			
FT SIGNAL			
FT 1			
FT 54			
FT SIGNAL			
FT 54			
FT POTENTIAL.			
SO SEQUENCE			
SO 507 AA; 55730 MW; 2B2E0A5E958FE5A CRC64;			
Query Match	68.4%;	Score 54;	DB 10; Length 507;
Best Local Similarity	71.4%;	Pred. No. 0.2;	
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;			

Db 236 GIDIFASKRFFLEK 249

RESULT 2

Q9M6Y2 PRELIMINARY; PRT; 503 AA.
AC Q9M6Y2; 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Proteanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
NCBI_TaxID=8022;
OX NCB1_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADIPOSE TISSUE;
RA Lindberg A., Oliverson G.;
RT "Lipoprotein lipase from rainbow trout differs in several respects from the enzyme in mammals."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Prat F., Kwon J.Y., Tyler C.R.;
RT "Molecular characterization of putative yolk processing enzymes and their expression during oogenesis and embryogenesis in rainbow trout Oncorhynchus mykiss."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ224693; CAB40545.1; -
DR EMBL; AF358669; AAK69707.1; -
DR HSSP; P06857; IRP1.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR001024; Lipoxigenase_LH2.
DR InterPro; IPR000379; Ser-esterase_site.
DR Pfam; PF00151; Lipase; 1.
DR Pfam; PF01477; PLAT; 1.
DR PRINTS; PR00821; TAGLIPASE.
DR SMART; SM00308; LH2; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 503 AA; 57273 MW; FDB3BA08C9D0CA66 CRC64;
Query Match
Best Local Similarity 57.0%; Score 45; DB 13; Length 503;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 3 DIFASKRFFLEK 14
Db 429 DIFASKRFFLEK 440
RESULT 3
Q9AHX1 PRELIMINARY; PRT; 252 AA.
AC Q9AHX1; 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
OS ATP synthase gamma subunit.
GN ATPG.
OS Carsonella ruddii.
OC Bacteria; Proteobacteria; gamma subdivision; Candidatus Carsonella.
OX NCB1_TaxID=114186;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=2115546; PubMed=11222582;
RA Clark M.A., Baumann L., Thao M.L., Moran N.A., Baumann P.;
RT "Degenerative Minimalism in the genome of a Psyllid Endosymbiont."
J. Bacteriol. 183:1853-1861 (2001).
EMBL; AF291051; AAK17110.1; -

DR InterPro; IPR00131; ATPase_gamma.

DR Pfam; PF00231; ATP-Synt_1.
DR PRINTS; PR00126; ATPASEGAMMA.
SQ SEQUENCE 252 AA; 29902 MW; 450A3F71C8039A8D CRC64;

Query Match
Best Local Similarity 54.4%; Score 43; DB 2; Length 252;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 GIDIFASKRFFLEK 14
Db 106 GIDIFASKRFFLEK 119

RESULT 4

Q15807 PRELIMINARY; PRT; 1360 AA.
AC Q15807; 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
OS ATP-binding cassette protein 1 (Fragment).
GN ABC1.
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
NCBI_TaxID=5807;
OX NCB1_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KSU-1;
RA MEDLINE=97376910; PubMed=9233681;
RX Perkins M.E., Volkmann S., Wirth D.F., Le Blancq S.M.;
RT "Characterization of an ATP-binding cassette transporter in Cryptosporidium parvum."
J. Mol. Biochem. Parasitol. 87:117-122 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KSU-1;
RA Zapata F., Riojas Y.A., Wu T.W., Le Blancq S.M.;
RT "ABC proteins in Cryptosporidium parvum."
Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KSU-1;
RA Leblancq S.M., Riojas Y.A., Wu T.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90628; AAC47653.2; -
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABCtransporter.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 2.
DR PRODOM; PD000006; ABC_transporter; 3.
DR SMART; SM00382; AAA_2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding.
FT NON TER 1360 1360
SQ SEQUENCE 1360 AA; 154147 MW; 4ECD0DAD43521994 CRC64;
Query Match
Best Local Similarity 54.4%; Score 43; DB 5; Length 1360;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 3 DIFASKRFFLEK 15
Db 453 DIFASKRFFLEK 465
RESULT 5
Q9XYH6 PRELIMINARY; PRT; 1431 AA.
AC Q9XYH6; 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (Tremblrel. 12, last sequence update)
 DE 01-JUN-2002 (Tremblrel. 21, last annotation update)
 GN CPABC.
 OS Cryptosporidium parvum.
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 CC Cryptosporidiidae; Cryptosporidium.
 OX NCBI_TaxID=5807;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=SF6H1;
 RX MEDLINE=99254110; PubMed=10318953;
 RA Perkins M.E., Riojas V.A., Wu T.W., Le Blancq S.M.;
 RT "OpABC, a cryptosporidium parvum ATP-binding cassette protein at the
 host-parasite boundary in intracellular stages."
 RL Proc. Natl. Acad. Sci. U.S.A. 96:5734-5739 (1999).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL; AF110147; AAD32093.1; -
 DR HSSP; P13569; INBD.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001140; ABCtransporter.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transport; 2.
 DR SMART; SM00382; AAA_2
 DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Transport.
 KW
 SQ SEQUENCE 1431 AA; 162177 MW; 18ABAD49184EB4E CRC64;

Query Match 54.4%; Score 43; DB 5; Length 1431;
 Best Local Similarity 61.5%; Pred. No. 59;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 IDIFASKNFHL 15
 DB 453 DILSINFNIQKN 465

RESULT 6
 ID Q9PQ31 PRELIMINARY; PRT; 156 AA.
 AC Q9PQ31;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
 DE Hypothetical protein U0459.
 GN U0459.
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 CC Mycoplasmataceae; Ureaplasma.
 OX NCBI_TaxID=134821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SEROVAR 3;
 RX MEDLINE=20500219; PubMed=11048724;
 RA Glaes J.I., Iefkowitz E.J., Glaes J.S., Heiner C.R., Chen E.Y.;
 RT "The complete sequence of the mucosal pathogen Ureaplasma
 urealyticum."
 RL Nature 407:757-762 (2000).
 DR EMBL; AE002142; AAF30871.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 156 AA; 18839 MW; E50C5FA78412CA7 CRC64;

Query Match 53.3%; Score 42; DB 16; Length 156;
 Best Local Similarity 72.7%; Pred. No. 8.6;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 IDIFASKNFHL 12
 DB 66 IKIFKTNFHL 76

RESULT 7
 ID 045343 PRELIMINARY; PRT; 305 AA.
 AC 045343;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
 DE P10A3.12 protein.
 GN P10A3.12.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lloyd C.R.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none.

RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018 (1998).
 DR EMBL; Z92829; CAB07348.1; -
 DR InterPro; IPR003002; 7TM_chemol.
 DR InterPro; IPR000168; 7TM_nematode.
 DR Pfam; PF01461; 7tm_4; 1.
 KW
 SQ SEQUENCE 305 AA; 35272 MW; DB6CA7E6D6C1F4D CRC64;

Query Match 53.2%; Score 42; DB 5; Length 305;
 Best Local Similarity 72.7%; Pred. No. 18;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 IDIFASKNFHL 12
 DB 112 IEIFAKRFHL 122

RESULT 8
 ID Q6652 PRELIMINARY; PRT; 630 AA.
 AC Q6652;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE ORF 50.
 OS Equine herpesvirus type 2 (strain 86/87) (EHV-2).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Gammaherpesvirinae.
 OX NCBI_TaxID=82831;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95302501; PubMed=7783207;
 RA Telford E.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus 2."
 RL J. Mol. Biol. 249:520-528 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Telford E.A.R.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U20824; AAC13838.1; -
 DR InterPro; IPR004998; Herpes_TAF50.
 DR Pfam; PF03326; Herpes_TAF50; 1.
 KW
 SQ SEQUENCE 630 AA; 70500 MW; 1.93EBB281A7B060B5 CRC64;

Query Match 51.9%; Score 41; DB 12; Length 630;
 Best Local Similarity 50.0%; Pred. No. 57;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 IDIFASKNFHL 15
 DB 112 IEIFAKRFHL 122

Db 614 LDIISLHNLHURKS 627

RESULT 9

ID 081407 PRELIMINARY; PRT; 251 AA.
AC 081407;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE 1-aminocyclopropane-1-carboxylate synthase 4 (Fragment).
GN ACS4.
OS Sinapis arvensis (Charlock).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Sinapis.
OX NCBI_TaxID=29728;
RN [1]
RP SEQUENCE FROM N.A.
RA Wei Y.D., Hall J.C.;
RT "Characterization of the 1-aminocyclopropane-1-carboxylate synthase
gene family of Sinapis arvensis."
RI Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CORACOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
ONE AND A MITOCHONDRIAL ONE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
AMINOTRANSFERASES.
CC EMBL; AF074930; AAC32839.1; -.
DR HSSP; P37821; 1B86.
DR InterPro; IPR001176; ACC_synthase.
DR InterPro; IPR004839; Aminotransf1/2.
DR InterPro; IPR004838; Nitransf.1.
DR Pfam; PF00155; aminotran.1.2; 1.
DR PRINTS; PR00753; ACCSYNTASE.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KM Pyridoxal phosphate. 1
FT NON_TER 1 251
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28063 MW; DCD158F54CB993A CRC64;
Query Match 50.6%; Score 40; DB 10; Length 251;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 IDIFASKNFHL 12
Db 171 VDFIASKNFHL 181
RESULT 10
ID 08XJS9 PRELIMINARY; PRT; 445 AA.
AC 08XJS9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein CPE1675.
GN CPE1675.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RA RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohceni K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
flesh-eater."
Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).

DR EMBL; AP003191; BAB81381.1; -.
DR InterPro; IPR001861; UPF00004.
DR Pfam; PF00919; UPF00004; 1.
DR TIGRFAMs; TIGR01125; UPF00004; 1.
DR TIGRFAMs; TIGR00089; UPF00004; 1.
DR PROSITE; PS01278; UPF00004; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 445 AA; 50621 MW; 4C7720A60F18A390 CRC64;
Query Match 50.6%; Score 40; DB 16; Length 445;
Best Local Similarity 58.3%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHL 12
Db 207 GIDIFYGKMLHL 218
RESULT 11

ID 098PVL PRELIMINARY; PRT; 598 AA.
AC 098PVL;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Sugar ABC transporter permease protein.
GN MYPU 6180.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RA RC STRAIN=UAB CTIP.
RX MEDLINE=21267165; PubMed=11353084; Barde V., Samson D., Gällsson F.,
RA Chambaud I., Hellig R., Ferris S., Roeha E.P.C.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Roeha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
Mycoplasma pulmonis."
RT Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445565; CAC13791.1; -.
DR MyPulst; MYPU 6180; -.
DR InterPro; IPR001851; Bac_inmem_transp.
DR InterPro; IPR001064; Crystallin.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
KM Complete proteome.
SQ SEQUENCE 598 AA; 68868 MW; E030842FAC876DE5 CRC64;
Query Match 50.6%; Score 40; DB 16; Length 598;
Best Local Similarity 63.6%; Pred. No. 83;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHL 11
Db 383 GINIFASKNFHL 393
RESULT 12

ID 09VRL4 PRELIMINARY; PRT; 859 AA.
AC 09VRL4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE CG10633 protein.
GN CG10633.
OS Drosophila melanogaster (Fruit Fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Empidoidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]


```

Query Match Similarity      50.6%; Score 40; DB 5; Length 859;
Best Local Similarity      61.5%; Pred. No. 1.2e+02;
Matches      8; Conservative      2; Mismatches      3; Indels      0; Gaps      0;

QY      3      DIFASAKNPHLOKN 15
      | | | | | | | |
Db      81      DAFAAKNFOLIKS 93

RESULT 13
Q9XEY7
ID      Q9XEY7      PRELIMINARY;      PRT;      490 AA.
AC      Q9XEY7;
DT      01-NOV-1999 (TREMBLrel. 12, Created)
DT      01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Hypothetical 54.7 kDa protein.
OS      Oryza sativa (Rice).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Emarstioideae; Oryzaceae; Oryza.
NCBI_TaxID=4530;
OX      NCB1
RN
SEQUENCE FROM N.A.
RP
LA      Laca V., Lou A., Young S., Messing U.,
RT      "Microsynteny in a gene-dense region in maize, sorghum, and rice.";

```

RESULT 14			
087205			
ID	087205	PRELIMINARY;	PRT; 64 AA.
AC	087205;		
DT	01-NOV-1998 (TReMBLrel. 08, Created)		
DT	01-NOV-1998 (TReMBLrel. 08, Last sequence update)		
DT	01-NOV-1998 (TReMBLrel. 08, Last annotation update)		
DE	Hypothetical 8.2 kDa protein.		
OS	Lactococcus lactis.		
OC	Plasmid pMRC01.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;		
OC	Streptococcaceae; Lactococcus.		
OX	NCBI_TaxID=1358;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=DPC3147;		
RX	MEDLINE=99000510; Pubmed=9767571;		
RA	Dougherty B.A., Hill C., Weidman J.F., Richardson D.R., Venter J.C.,		
RA	Ross R.P.,		
RT	"Sequence and analysis of the 60 kb conjugative, bacteriocin-producing		
RT	plasmid pMRC01 from Lactococcus lactis DPC3147.",		
RL	Mol. Microbiol. 29:1029-1038(1998).		
DR	EMBL; AE001272; AAC56026.1; -		
KW	Hypothetical protein; Plasmid.		
SQ	SEQUENCE 64 AA; 8187 MW; 23713F5794FBE304 CRC64;		

Query Match	Best Local Similarity	Score	DB	Length
49.4%;	61.5%;	39;	2;	64;
Matches	8; Conservative	1;	Mismatches	4;
			Indels	0;
			Gaps	0;

Sat May 10 08:43:28 2003

us-09-142-524d-120.rspt

Page 6

RL Virus Genes 0:0-0(1997).
DR EMBL: U86902; AAB57951.1; -
FT NON TER 122 122
SQ SEQUENCE 122 AA; 14361 MW; 2284E1A5B77492F CRC64;

Query Match 49.4%; Score 39; DB 12; Length 122;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 DIFASKNFHL 12
|:|:|:|:|:
DB 54 DVFAVNFHV 63

Search completed: April 20, 2003, 13:13:16
Job time : 14.6711 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds
(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524d-121

Perfect score: 79

Sequence: 1 ASKNFHLQKNTIGTG 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

PIR 73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	514	2 S48730	Cry j II protein -
2	79	100.0	514	2 JC2498	second major aller
3	54	68.4	514	2 JC7100	polygalacturonase
4	51	64.6	507	2 JC7366	Jun a 2 protein -
5	45	57.0	419	2 T38199	conserved hypochet
6	45	57.0	513	2 T11686	aspartic proteinase
7	43	54.4	97	1 R6RT37	ribosomal protein
8	43	54.4	97	2 S47646	ribosomal protein
9	43	54.4	275	2 T28738	hypothetical prote
10	42	53.2	421	2 B83615	probable porin PA0
11	41	51.9	491	2 D97037	hypothetical prote
12	41	51.9	530	1 A27481	lanosterol 14alpha
13	40	50.6	386	2 E71693	hypothetical prote
14	40	50.6	549	1 S48474	glucan 1,4-alpha-g
15	40	50.6	767	1 JU0474	glucan 1,4-alpha-g
16	40	50.6	778	1 ALBYG	T-cell receptor al
17	39	49.4	104	2 S07705	T-cell receptor al
18	39	49.4	104	2 G24402	T-cell receptor al
19	39	49.4	109	2 F81669	conserved hypochet
20	39	49.4	115	2 C24891	T-cell receptor al
21	39	49.4	116	2 D24891	T-cell receptor al
22	39	49.4	118	2 I71934	MHC class II I-A-a
23	39	49.4	121	2 A29080	T-cell receptor al
24	39	49.4	133	2 F24402	T-cell receptor al
25	39	49.4	342	2 A95165	hypothetical prote
26	39	49.4	346	2 H98030	S-adenosylmethioni
27	39	49.4	352	2 G81921	hypothetical prote
28	39	49.4	391	2 T51609	polygalacturonase
29	39	49.4	723	2 H85092	hypothetical prote

ALIGNMENTS

30 39 49.4 874 2 A10281
31 38.5 48.7 819 2 T30050
32 38 48.1 60 2 A30606
33 38 48.1 173 2 JN0515
34 38 48.1 174 2 S71594
35 38 48.1 181 2 AF1493
36 38 48.1 250 2 AB1766
37 38 48.1 250 2 A11366
38 38 48.1 262 2 C97426
39 38 48.1 262 2 AC2644
40 38 48.1 324 2 AG2956
41 38 48.1 324 2 G98326
42 38 48.1 341 1 DEECTH
43 38 48.1 341 2 F91190
44 38 48.1 341 2 G86037
45 38 48.1 343 2 B82405

probable insectici
hypothetical prote
T-cell receptor al
FliV protein - Sal
pathogenesis-relat
hypothetical prote
regulatory protein
regulatory protein
flagellar basal-bo
flagellar basal-bo
hypothetical prote
c4-dicarboxylate-b
L-threonine 3-dehy
threonine dehydrog
threonine dehydrog
threonine 3-dehydr

RESULT 1
S48730
Cry j II protein - Japanese cedar
C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C/Accession: S48730
R/Namba, M.; Kurose, M.; Torioka, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.;
FEMS Lett. 353, 124-128, 1994
A/Title: Molecular cloning of the second major allergen, Cry j II, from Japanese ceda
A/Reference number: S48730; MWID:95010777; PMID:7926035
A/Accession: S48730
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-514 <NAM>
A/Cross-References: GB:D37765; NID:9577695; PIDN:BA07021.1; PID:dl007598; PID:g577694

Query Match 100.0%; Score 79; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ASKNFHLQKNTIGTG 15
Db 240 ASKNFHLQKNTIGTG 254

RESULT 2
JC2498
second major allergen Cry j II precursor - Japanese cedar
C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C/Accession: JC2498; PC2346; A60147
R/Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
A/Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese
A/Reference number: JC2498; MWID:94271186; PMID:8002972
A/Accession: JC2498
A/Molecule type: mRNA
A/Residues: 1-514 <KOM>
A/Cross-References: DDBJ:D29772; NID:9506857; PIDN:BA06172.1; PID:g506858
A/Accession: PC2346
A/Molecule type: protein
A/Residues: 52-61 <KO2>
R/Sakaguchi, M.; Inouye, S.; Tanitai, M.; Ando, S.; Usui, M.; Matuhasi, T.
Allergy 45, 309-312, 1990
A/Title: Identification of the second major allergen of Japanese cedar pollen.
A/Reference number: A60147; MWID:90342988; PMID:2382797
A/Accession: A60147
A/Molecule type: protein
A/Residues: 55-64 <SAK>
C/Keywords: glycoprotein; pollen
F1-54/Domain: signal sequence #status predicted <SIG>
F155-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match
Best Local Similarity 100.0%; Score 79; DB 2; Length 514;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASKNPHLQKNTTGTG 15
DB 240 ASKNPHLQKNTTGTG 254

RESULT 3

JC7100 polygalacturonase Cha o 2 - Japanese cypress

C/Species: Chamaecyparis obtusa (Japanese cypress)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C/Accession: JC7100; PC7026
R/Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.
Biochem. Biophys. Res. Commun. 263, 166-171, 1999
A/Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all
A/Reference number: JC7100; MUID:99417540; PMID:10486272
A/Accession: JC7100
A/Molecule type: mRNA
A/Residues: 1-514 <MOR>
A/Accession: PC7026
A/Molecule type: protein
A/Residues: 51-62 <MO2>

Query Match
Best Local Similarity 66.7%; Score 54; DB 2; Length 514;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ASKNPHLQKNTTGTG 15
DB 240 ASKNPHLQKNTTGTG 254

RESULT 4

JC7366 Jun a 2 protein - mountain cedar

C/Species: Juniperus ashei (mountain cedar)
C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C/Accession: JC7366; PC7093
R/Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.
Biochem. Biophys. Res. Commun. 275, 195-202, 2000
A/Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all
A/Reference number: JC7366
A/Accession: JC7366
A/Molecule type: mRNA
A/Residues: 1-507 <YOK>
A/Cross-references: GB:AJ04653
A/Accession: PC7093
A/Molecule type: protein
A/Residues: 55-63 <Y02>
C/Comment: This protein, a second major allergen of mountain cedar pollen, which is inv
C/Keywords: glycoprotein; pollen

Query Match
Best Local Similarity 64.6%; Score 51; DB 2; Length 507;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ASKNPHLQKNTTGTG 15
DB 241 ASKNPHLQKNTTGTG 255

RESULT 5

T38199 conserved hypothetical protein SPAC22F8.09 - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T38199

R/Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.; Ramsperger, U.; Pohl, T.
submitted to the EMBL Data Library, August 1999

A/Reference number: Z21777
A/Accession: T38199
A/Status: preliminary; translated from GB/EMBL/DDBU
A/Molecule type: DNA
A/Residues: 1-419 <LYN>
A/Cross-references: EMBL:AL109811; PIR:GAB52719.1; GSPDB:GN00066; SPDB:SPAC22F8.09
A/Experimental source: strain 972h-, cosmid c22F8
C/Genetics:
A/Map position: 1
A/Introns: 46/2

Query Match
Best Local Similarity 57.0%; Score 45; DB 2; Length 419;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 KNPHLQKNTT 12
DB 410 KDPHLQKNSI 419

RESULT 6

T11686 aspartic proteinase (EC 3.4.23.-) - cowpea

C/Species: Vigna unguiculata (cowpea)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C/Accession: T11686
R/d'Arcy-Lamer, A.; Zully-Fodil, Y.; Pham Thi, A.T.; Ferrarri-Ilhou, R.
submitted to the EMBL Data Library, June 1996
A/Description: Molecular cloning and nucleotide sequence of a cDNA encoding an aspartic
A/Reference number: Z17314
A/Accession: T11686
A/Status: preliminary; translated from GB/EMBL/DDBU
A/Molecule type: mRNA
A/Residues: 1-513 <DAR>
A/Cross-references: EMBL:U61396; NID:G1420935; PID:G1420936
A/Experimental source: strain 1183
C/Superfamily: oryzasin; saposin repeat homology
C/Keywords: aspartic proteinase; hydrolase

Query Match
Best Local Similarity 57.0%; Score 45; DB 2; Length 513;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 KNPHLQKNTTGTG 15
DB 62 KPHLQNNFAGTG 74

RESULT 7

R6RT37 ribosomal protein L37, cytosolic [validated] - rat

C/Species: Rattus norvegicus (Norway rat)
C/Date: 25-Feb-1995 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C/Accession: JN0478; JN0455; S21496; A02779; S11421
R/Chan, Y.L.; Paz, V.; Olivera, J.; Wool, I.G.
Biochem. Biophys. Res. Commun. 192, 590-596, 1993
A/Title: The primary structure of L37 - a rat ribosomal protein with a zinc finger-1
A/Reference number: JN0478; MUID:93249430; PMID:8484768
A/Accession: JN0478

A/Accession: JN0478
A/Molecule type: mRNA
A/Residues: 1-97 <CHA>
A/Accession: PNO455
A/Molecule type: protein
A/Residues: 2-40 <CH1>

A/Note: The protein is designated as ribosomal protein L37
R/Chan, Y.L.; Paz, V.; Olivera, J.; Wool, I.G.
submitted to the EMBL Data Library, May 1992
A/Description: The primary structure of rat ribosomal protein S9.
A/Reference number: S21496
A/Accession: S21496

A/Molecule type: mRNA
 A/Residues: 1-95 <CHW>
 A/Cross-references: EMBL:X66369
 R/Lin, A.; McNally, J.; Wool, I.G.
 J. Biol. Chem. 258, 10664-10671, 1983
 A/Title: The primary structure of rat liver ribosomal protein L37. Homology with yeast
 A/Reference number: A02779; MUID:83291000; PMID:6350292
 A/Accession: A02779
 A/Molecule type: protein
 A/Residues: 'N', 3-18, 'G', 20-21, 'KSKGKALOKGNSFPKLT', 22-25, 26-30, 'SKKYNTLHAKRTTGM', 40-4
 A/Note: the protein is designated as ribosomal protein L37
 R/Wittmann-Liebold, B.; Geisler, A.W.; Lin, A.; Wool, I.G.
 J. Supramol. Struct. 12, 425-433, 1979
 A/Title: Sequence of the amino-terminal region of rat liver ribosomal proteins S4, S6, S
 A/Reference number: S11413; MUID:80252792; PMID:398910
 A/Accession: S11421
 A/Molecule type: protein
 A/Residues: 'X', 3-18, 'G', 20-21, 'KKXGKA', 29-31 <MT>
 A/Note: the protein is designated as ribosomal protein L37
 C/Superfamily: rat ribosomal protein L37
 C/Keywords: cytosol; protein biosynthesis; ribosome; zinc finger
 F/2-97/Product: ribosomal protein L37 #status experimental <RPL>
 F/19-37/Region: zinc finger CCCC motif

Query Match 54.4%; Score 43; DB 1; Length 97;
 Best Local Similarity 66.7%; Pred. No. 1;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SKNPHLOKNTIG 13
 Db 24 SKNPHLOKSTCG 35

RESULT 8
 S4746
 ribosomal protein L37, cytosolic - human
 C/Species: Homo sapiens (man)
 C/Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 13-Aug-1999
 C/Accession: S47646; S68344
 R/Barnard, G.F.; Staliunas, R.J.; Puder, M.; Steele Jr., G.D.; Chen, L.B.
 Biochim. Biophys. Acta 1218, 425-428, 1994
 A/Title: Human ribosomal protein L37 has motifs predicting serine/threonine phosphorylat
 A/Reference number: S47646; MUID:94325352; PMID:7545944
 A/Accession: S47646
 A/Molecule type: mRNA
 A/Residues: 1-97 <BAR>
 A/Cross-references: EMBL:L11567; NID:G292440; PIDN:AA62148.1; PID:G292441
 R/Su, S.; Bird, R.C.
 Eur. J. Biochem. 232, 789-797, 1995
 A/Title: Cell cycle, differentiation and tissue-independent expression of ribosomal prot
 A/Reference number: S68344; MUID:96028120; PMID:7588717
 A/Accession: S68344
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-97 <SUS>
 A/Cross-references: GB:S79979; NID:G1839333; PIDN:AA647039.1; PID:G1839334
 C/Genetics: GDB:RPL37
 A/Map position: 14pter-14qter
 C/Superfamily: rat ribosomal protein L37
 C/Keyword: protein biosynthesis; ribosome; zinc finger

Query Match 54.4%; Score 43; DB 2; Length 97;
 Best Local Similarity 66.7%; Pred. No. 1;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SKNPHLOKNTIG 13
 Db 24 SKNPHLOKSTCG 35

RESULT 9

T28738
 hypothetical protein F48G7.4 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T28738
 R/Clarke, K.; Wohldmann, P.; Harrison, M.
 submitted to the EMBL Data Library, January 1998
 A/Description: The sequence of C. elegans cosmid F48G7.
 A/Reference number: Z20517
 A/Accession: T28738
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-275 <CIA>
 A/Cross-references: EMBL:AF039044; PIDN:AA647944.1; GSPDB:GN00023; CESP:F48G7.4
 A/Experimental source: strain Bristol NZ; clone F48G7
 C/Genetics:
 A/Gene: CESP:F48G7.4
 A/Map position: 5
 A/Introns: 189/3; 232/2

Query Match 54.4%; Score 43; DB 2; Length 275;
 Best Local Similarity 64.3%; Pred. No. 3.4;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 SKNPHLOKNTIGTG 15
 Db 230 SKRFTYQKQIGTG 243

RESULT 10
 B83615
 probable porin PA0240 [imported] - Pseudomonas aeruginosa (strain PA01)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Jun-2001
 C/Accession: B83615
 R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
 ; Loy, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat
 A/Reference number: B83615
 A/Accession: B83615
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-421 <STO>
 A/Cross-references: GB:AE004462; GB:AE004091; NID:G9946077; PIDN:AA603629.1; GSPDB:GN
 A/Experimental source: strain PA01
 C/Genetics:
 A/Gene: PA0240
 C/Superfamily: Pseudomonas aeruginosa outer membrane porin PA0958

Query Match 53.2%; Score 42; DB 2; Length 421;
 Best Local Similarity 50.0%; Pred. No. 8.5;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 SKNPHLOKNTIG 13
 Db 43 ARNPHLRNFTVG 54

RESULT 11
 D97037
 hypothetical protein CAC113 [imported] - Clostridium acetobutylicum
 C/Species: Clostridium acetobutylicum
 C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C/Accession: D97037
 R/Nolling, U.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Le
 ; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C
 A/Reference number: A96900; MUID:21359325; PMID:21359325
 A/Accession: D97037
 A/Status: preliminary

A:Molecule type: DNA
A:Residues: 1-491 <KOR>
A:Cross-references: GB:AE001437; PIDN:AAK79087.1; PID:g15024032; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC113

Query Match 51.9%; Score 41; DB 2; Length 491;
Best Local Similarity 57.1%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SKNFHOKNTIGTG 15
DB 326 SWNFHINTIGTG 339

RESULT 12

A27491
lanosterol 14alpha-demethylase (EC 1.14.14.-) cytochrome P450 51 - yeast (Saccharomyces

N:Alternate names: protein YHR07C
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 23-Mar-2001
C:Accession: A27491; S46804; B31563; A25563

R:Kahl, V.F.; Woods, C.W.; Turf, T.G.; Dey, C.R.; Sutter, T.R.; Loper, J.C.
DNA 6, 529-537, 1987

A:Title: Primary structure of the P450 lanosterol demethylase gene from Saccharomyces ce
A:Reference number: A27491; MUID:88111027; PMID:3322742
A:Accession: A27491

A:Molecule type: DNA
A:Residues: 1-530 <KAL1>
A:Cross-references: EMBL:M18109; NID:g170945; PIDN:AAA34379.1; PID:g170946

R:Favell, T.
Submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9780.

A:Reference number: S46797
A:Accession: S46804

A:Molecule type: DNA
A:Residues: 1-530 <PAV>

A:Cross-references: EMBL:U10555; NID:g500813; PIDN:AA68433.1; PID:g500824; GSPDB:GN0000
R:Ikshida, N.; Aoyama, Y.; Hatanaka, R.; Oyama, Y.; Inajo, S.; Ishiguro, M.; Oshima, T.;
Biochem. Biophys. Res. Commun. 155, 317-323, 1988

A:Title: A single amino acid substitution converts cytochrome P450-14DM to an inactive fi
A:Reference number: A31569; MUID:88326319; PMID:3046615
A:Accession: B31569

A:Molecule type: DNA
A:Residues: 1-432, 'N', 434-530 <ISH>
A:Cross-references: EMBL:M21403; NID:g171353; PIDN:AAA34546.1; PID:g171354

R:Kahl, V.F.; Loper, J.C.; Dey, C.R.; Woods, C.W.; Sutter, T.R.
Gene 45, 237-245, 1986
A:Title: Isolation of a cytochrome P-450 structural gene from Saccharomyces cerevisiae.

A:Reference number: A25563; MUID:87106820; PMID:3542713
A:Accession: A25563

A:Molecule type: DNA
A:Residues: 444-530 <KAL2>
A:Cross-references: EMBL:M15663

A:Genetics: SGD:ERG11; MIPS:YHR007C
A:Cross-references: SGD:S0001049; MIPS:YHR007C

A:Map position: 8R
C:Superfamily: human cytochrome P450 CYP51; cytochrome P450 homolog
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;
F:311492/Domain: cytochrome P450 homolog <P45>
F:470/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 51.9%; Score 41; DB 1; Length 530;
Best Local Similarity 61.5%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 SKNFHOKNTIGTG 14
DB 181 SKNFRLNRTIGTG 193

RESULT 13

E71693
hypothetical protein RP364 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii
C>Date: 21-Nov-1998 #sequence revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: E71693

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark,
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: E71693

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-386 <AND>
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14823.1; PID:g386
A:Experimental source: strain Madrid E
C:Genetics:

A:Gene: RP364

Query Match 50.6%; Score 40; DB 2; Length 386;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 FHLOKNTIGTG 15
DB 214 FHLOKNTIGTG 224

RESULT 14

S48474
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) SGAL - yeast (Saccharomyces cerevisiae)

N:Alternate names: glucanase; protein YIL099W
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 12-Nov-1999

C:Accession: S48474; C26877; S27284
R:Bowman, S.; Churche, C.
Submitted to the EMBL Data Library, September 1994

A:Reference number: S48455
A:Accession: S48474

A:Molecule type: DNA
A:Residues: 1-549 <BDW>

A:Cross-references: GB:247047; EMBL:Z38125; NID:g603997; PID:g763247; GSPDB:GN00009;
R:Yamashita, I.; Nakamura, M.; Fukui, S.
J. Bacteriol. 169, 2142-2149, 1987

A:Title: Gene fusion is a possible mechanism underlying the evolution of STAL.
A:Reference number: A91831; MUID:87194600; PMID:3106330
A:Accession: C26877

A:Molecule type: DNA
A:Residues: 1-503,507, 'W', 513-514, 'TG', 516 <YAM>
A:Cross-references: EMBL:M16166; NID:g172592; PIDN:AAA35042.1; PID:g172593

R:Pardo, J.M.; Ianez, E.; Zalacain, M.; Claros, M.G.; Jimenez, A.
FEBS Lett. 239, 179-184, 1988
A:Title: Similar short elements in the 5' regions of the STAL and SGA genes from Sacc

A:Reference number: S27281; MUID:89031230; PMID:3141213
A:Accession: S27284

A:Molecule type: DNA
A:Residues: 1-183, 'H', 185-190 <PAR>
A:Cross-references: EMBL:X13858; NID:g4461; PIDN:CAA32071.1; PID:g4463

C:Genetics: SGD:SGA1; MIPS:YIL099W
A:Cross-references: SGD:S0001361; MIPS:YIL099W

A:Map position: 9L
C:Function:
A:Description: hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively
C:Superfamily: yeast glucan 1,4-alpha-glucosidase SGAL; glucan 1,4-alpha-glucosidase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation; yeast vacuole
F:77-536/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match 50.6%; Score 40; DB 1; Length 549;
Best Local Similarity 57.1%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 4 NFHLOK-NTIGTG 15

Db 159 SFHLQSRNNTLTGAG 172

RESULT 15

JU0474

glucan 1,4-alpha-glucosidase (EC 3.2.1.3) STA2 - yeast (*Saccharomyces diastaticus*)

N/Alternate names: glucoamylase II

C/Species: *Saccharomyces diastaticus*

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: JU0474; S35895

R/Lambrechts, M.G.; Pretorius, I.S.; Sollitt, P.; Marmur, J.

Gene 100, 95-103, 1991

A/Title: Primary structure and regulation of a glucoamylase-encoding gene (STA2) in *Sac*

A/Reference number: JU0474; MUID:91276266; PMID:2055484

A/Accession: JU0474

A/Molecule type: DNA

A/Residues: 1-767 <LAM>

A/Cross-references: GB:M60650; NID:g172733; PIDN:AAA35107.1; PID:g172734

A/Note: the authors translated the codon CCG for residue 337 as Ala, CCG for residue 364

R/Kim, K.; Bajszar, G.; Lee, S.Y.; Knudsen, F.; Mattoon, J.R.

A/Description: Cloning of a new allelic variant of a *Saccharomyces diastaticus* glucoamyl

A/Reference number: S35895

A/Accession: S35895

A/Molecule type: DNA

A/Residues: 1-163, 'A', 164-622, 'D', 624-767 <KIM>

A/Cross-references: EMBL:M90490; NID:g172735; PIDN:AAA20560.1; PID:g172736

C/Genetics:

A/Gene: STA2; DEX1

C/Superfamily: yeast glucan 1,4-alpha-glucosidase STA1; glucan 1,4-alpha-glucosidase hom

C/Keywords: glycosidase; hydrolase; polysaccharide degradation

F/334-767/Domain: glucan 1,4-alpha-glucosidase homology <GAG>

Query Match 50.6%; Score 40; DB 1; Length 767;

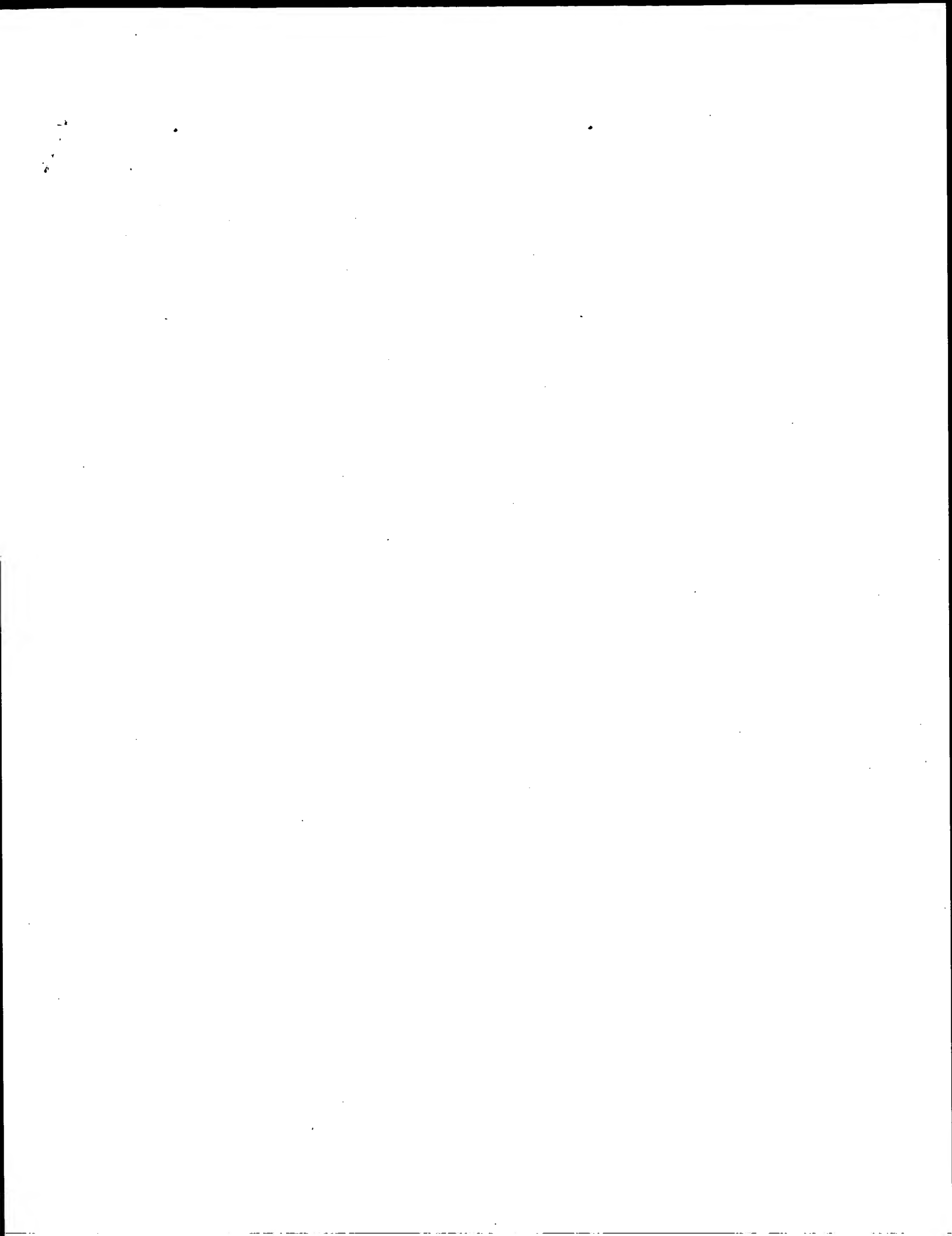
Best Local Similarity 57.1%; Pred. No. 40;

Matches 8; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 4 NFHLQK--NTIGTG 15

Db 416 SFHLQRTNTLTGAG 429

Search completed: April 20, 2003, 13:15:55
 Job time: 7.07895 secs



GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524d-121

Perfect score: 79

Sequence: 1 ASKNFHLQKXITGTG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	514	1	MPA2_CRYVA
2	45	57.0	419	1	YIV9_SCHPO
3	43	54.4	92	1	RU37_BOVIN
4	43	54.4	96	1	RU37_HUMAN
5	43	54.4	96	1	RL37_ICTPU
6	41	51.9	530	1	CP51_YEAST
7	40	50.6	385	1	Y364_RICPR
8	40	50.6	549	1	AMYG_YEAST
9	40	50.6	767	1	AMYH_YEAST
10	40	50.6	768	1	AMYH_SACDI
11	39	49.4	234	1	A29B_DROME
12	39	49.4	342	1	QUEA_STREY
13	39	49.4	342	1	QUEA_STREY
14	39	49.4	437	1	PRPG_LACDL
15	38	48.1	173	1	FLIV_SALMU
16	38	48.1	262	1	FLUG_AGRIS
17	38	48.1	341	1	TDH_ECOLI
18	38	48.1	401	1	FLIB_SALCH
19	38	48.1	401	1	FLIB_SALCH
20	37	46.8	64	1	RL35_HELPY
21	37	46.8	262	1	FLUG_RHIME
22	37	46.8	371	1	LE3B_ASPEG
23	37	46.8	380	1	S18C_HUMAN
24	37	46.8	571	1	EXOI_SCHPO
25	37	46.8	657	1	IYRB_MYCPN
26	37	46.8	1034	1	GCSP_YEAST
27	37	46.8	1305	1	RKPL_YEAST
28	36.5	46.2	416	1	PNXI_XENLA
29	36.5	46.2	608	1	DEXT_PENNI
30	36	45.6	85	1	U186_DROME
31	36	45.6	86	1	U186_HUMAN
32	36	45.6	86	1	U186_MOUSE
33	36	45.6	132	1	TVA3_MOUSE

34	36	45.6	200	1	VATE_BORBU	05123 borrelia bu
35	36	45.6	380	1	Y103_MYCPN	P75530 mycoplasma
36	36	45.6	280	1	CYPE_DROME	Q9V363 drosophila
37	36	45.6	341	1	QUEA_CLOAB	Q9V362 drosophila
38	36	45.6	341	1	QUEA_STRAAM	Q9V361 campylobact
39	36	45.6	342	1	QUEA_CAMJE	Q92b13 listeria in
40	36	45.6	342	1	QUEA_LISMO	08Y629 rhizobium m
41	36	45.6	342	1	QUEA_LISMO	052998 rhizobium m
42	36	45.6	344	1	TDH_RHIME	Q9C166 lactococcus
43	36	45.6	345	1	QUEA_LACLA	Q9JW05 neisseria m
44	36	45.6	346	1	QUEA_NEIMA	09JW44 neisseria m
45	36	45.6	346	1	QUEA_NEIMA	

ALIGNMENTS

```

RESULT 1
MPA2_CRYVA
ID MPA2_CRYVA STANDARD: PRT; 514 AA.
AC P43212;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)
DE (Major pollen allergen Cry j 2) (Cry j II).
OS Cryptomeria japonica (Japanese cedar).
OC Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OX NCBI_TaxID=3369;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=95010777; PubMed=7926035;
RA Namba M., Kurose M., Totigoe K., Hino K., Taniguchi Y., Fukuda S.,
RA Usui M., Kurimoto M.;
RT "Molecular cloning of the second major allergen, Cry j II, from
RT Japanese cedar pollen."
RL FEBS Lett. 353:124-128(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=94271186; PubMed=8002972;
RA Komiyama N., Some T., Shimizu K., Morikubo K., Kino K.;
RT "cDNA cloning and expression of Cry j II the second major allergen of
RT Japanese cedar pollen."
RL Biochem. Biophys. Res. Commun. 201:1021-1028(1994).
RN [3]
RP SEQUENCE OF 55-64.
RX MEDLINE=90342988; PubMed=2382797;
RA Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matuhashi T.;
RT "Identification of the second major allergen of Japanese cedar
RT pollen."
RL Allergy 45:309-312(1990).
-1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
galactosiduronic linkages in pectate and other galacturonans.
-1- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).
-1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
(POLYGALACTURONASES).
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: D37765; BAA07021.1;
EMBL: D29772; BAA06172.1;
HSSP: P26509; 1BHE.
InterPro: IPR000743; GH28.
Pfam: PF00295; Glyco_hydro_28; 1.

```

DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 KM Hydrolase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
 KW Amyloplast; Glycoprotein; Allergen.
 FT SIGNAL 1 22
 FT PROPEP 23 45
 FT CHAIN 46 433
 FT PROPEP 434 514
 FT ACT SITE 278 278
 FT CARBOHYD 460 460
 FT CARBOHYD 472 472
 FT CARBOHYD 5 5
 FT CONFLICT 12 12
 FT CONFLICT 34 35
 FT CONFLICT 37 37
 FT CONFLICT 88 88
 FT CONFLICT 98 98
 FT CONFLICT 451 451
 FT CONFLICT 454 454
 FT CONFLICT 504 504
 FT CONFLICT 507 507
 SQ SEQUENCE 514 AA; 56645 MW; 624611C3FAD6302 CRC64;
 Query Match 100.0%; Score 79; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASKNFHLQKNTTGTG 15
 DB 240 ASKNFHLQKNTTGTG 254
 RESULT 2
 ID Y1Y9 SCHPO STANDARD; PRT; 419 AA.
 AC Q9UUT4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C22F8.09 in chromosome 1.
 GN SPAC22F8.09.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moutie S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckert G., Aert R., Robben J., Grymopiez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Cabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Moutier S.,
 RA Galibert F., Ayes S.J., Xiang Z., Hunt C., Moore K., Huest S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Moreno S., Armstrong J., Forsburg S.L.,
 RA Dominguez A., Revuelta J.U., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,

RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO THE GITSR2 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; ALI09831; CAB52719.1; -
 DR Hypothetical protein.
 SQ SEQUENCE 419 AA; 48129 MW; 9C2A72BE39E61DNA CRC64;
 Query Match 57.0%; Score 45; DB 1; Length 419;
 Best Local Similarity 80.0%; Pred. No. 0.77;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 3 KNFHLQKNTI 12
 DB 410 KNFHLQKNTI 419
 RESULT 3
 ID RL37 BOVIN STANDARD; PRT; 92 AA.
 AC P79244;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE 60S ribosomal protein L37 (Fragment).
 GN RPL37.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 NX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96028120; PubMed=7588717;
 RA Su S., Bird R.C.;
 RT "Cell cycle, differentiation and tissue-independent expression of
 RT ribosomal protein L37.";
 RL Eur. J. Biochem. 232:789-797(1995).
 CC -1- SIMILARITY: BELONGS TO THE L37E FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; S79980; AAD14319.1; -
 DR InterPro: IPR001569; Ribosomal L37E.
 DR Pfam: PF01907; Ribosomal_L37E; 1.
 DR ProDom: PD005132; Ribosomal_L37E; 1.
 DR PROSITE; PS01077; RIBOSOMAL_L37E; PARTIAL.
 KW Ribosomal protein.
 FT INIT MET 0
 FT NON TER 92
 FT SEQUENCE 92 AA; 10576 MW; 881F2879AA8E0CF8 CRC64;
 Query Match 54.4%; Score 43; DB 1; Length 92;
 Best Local Similarity 66.7%; Pred. No. 0.34;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 SKNFHLQKNTTGTG 13
 DB 23 SKNFHLQKNTTGTG 34

RESULT 4
ID RL37_HUMAN STANDARD; PRT; 96 AA.
AC P02403; O99883;
DT 21-JUN-1986 (Rel. 01, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L37 (G1.16).
GN RPL37.
OS Homo sapiens (Human), and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 10116;
RN [1]
RP SEQUENCE FROM N.A.
RA SPECIES=Human; TISSUE=Lymphoma;
RA Kato S.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA SPECIES=Human; TISSUE=Colon;
RX MEDLINE=94325352; PubMed=7545944;
RA Barnard G.F., Stanlunas R.J., Puder M., Steele G.D. Jr., Chen L.B.;
RT "Human ribosomal protein L37 has motifs predicting serine/threonine
RL phosphorylation and a zinc-finger domain.";
RL Biochim. Biophys. Acta 1218:425-428(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA SPECIES=Human, and Rat;
RX MEDLINE=96028120; PubMed=7588717;
RA Su S., Bird R.C.;
RT "Cell cycle, differentiation and tissue-independent expression of
RL ribosomal protein L37.";
RL Eur. J. Biochem. 232:789-797(1995).
RN [4]
RP SEQUENCE FROM N.A.
RA SPECIES=Human;
RX MEDLINE=21864036; PubMed=11875025;
RA Yoshizama M., Uechi T., Asakawa S., Kawasaki K., Kato S., Higa S.,
RA Maeda N., Minoshima S., Tanaka T., Shimizu N., Kemmochi N.;
RT "The human ribosomal protein genes: sequencing and comparative
RL analysis of 73 genes.";
RL Genome Res. 12:379-390(2002).
RN [5]
RP SEQUENCE OF 47-73 AND 75-96 FROM N.A.
RA SPECIES=Human;
RX MEDLINE=98248690; PubMed=9582194;
RA Kemmochi N., Kawaguchi T., Rozen S., Davis E., Goodman N.,
RA Hudson T.J., Tanaka T., Page D.C.;
RT "A map of 75 human ribosomal protein genes.";
RL Genome Res. 8:509-523(1998).
RN [6]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-39.
RA SPECIES=Rat; STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=93249430; PubMed=8484768;
RA Chan Y.-L., Paz V., Olivera J., Wool I.G.;
RT "The primary structure of L37 -- a rat ribosomal protein with a zinc
RL finger-like motif.";
RL Biochem. Biophys. Res. Commun. 192:590-596(1993).
RN [7]
RP PRELIMINARY SEQUENCE.
RA SPECIES=Rat;
RX MEDLINE=83291000; PubMed=6350292;
RA Lin A., McNally J., Wool I.G.;
RT "The primary structure of rat liver ribosomal protein L37. Homology
RL with yeast and bacterial ribosomal proteins.";
RL J. Biol. Chem. 258:10664-10671(1983).
CC -1- SIMILARITY: BELONGS TO THE L37E FAMILY OF RIBOSOMAL PROTEINS.
CC -1- CAUTION: REF. 7 SEQUENCE DIFFERS FROM THAT SHOWN EXTENSIVELY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sb-sib.ch).

DR EMBL; D23661; BAA0488.1; -;
DR EMBL; L11567; AAB62148.1; -;
DR EMBL; S79979; AAB47039.1; -;
DR EMBL; S79981; -; NOT ANNOTATED_CDS.
DR EMBL; AB061834; BAB79472.1; -;
DR EMBL; AB007184; BAA25843.1; -;
DR EMBL; AB007183; BAA25842.1; -;
DR EMBL; X63659; CAA47012.1; -;
DR PIR; A02779; R6RT37.
DR PIR; JN0478; JN0478.
DR PIR; S21496; S21496.
DR Genew; HGNC:10347; RPL37.
DR MIM; 604181; -;
DR InterPro; IPR001569; Ribosomal_L37E.
DR Pfam; PF01907; Ribosomal_L37e; 1.
DR ProDom; PD005132; Ribosomal_L37E; 1.
DR PROSITE; PS01077; RIBOSOMAL_L37E; 1.
DR KX Ribosomal protein.
FT INIT MET 0
FT CONFICT 0
SQ SEQUENCE 96 AA; 10947 MW; 959C318791ADE39 CRC64;
QY 2 SKNFHLOKNTIG 13
Db 23 SKAYHLQKSTCG 34

Query Match 54.4%; Score 43; DB 1; Length 96;
Best Local Similarity 66.7%; Pred. No. 0.35;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 5
ID RL37 ICTPU STANDARD; PRT; 96 AA.
AC Q90YU1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L37.
GN RPL37.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RP SEQUENCE FROM N.A.
RA Paterson A.P., Karai A., Liu Z.J.;
RT "Translational machinery of channel catfish: II. Complementary DNA and
RL expression of the complete set of 47 60S ribosomal proteins.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L37E FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sb-sib.ch).

DR EMBL; AF401593; AAK95165.1; -;
DR InterPro; IPR001569; Ribosomal_L37E.
DR Pfam; PF01907; Ribosomal_L37e; 1.
DR ProDom; PD005132; Ribosomal_L37E; 1.

DR PROSITE; PS01077; RIBOSOMAL_L37E; 1.
 KM Ribosomal protein.
 FT INT MET 0 BY SIMILARITY.
 SQ SEQUENCE 96 AA; 10911 MW; 94A618F26055E983 CRC64;
 Query Match 54.4%; Score 43; DB 1; Length 96;
 Best Local Similarity 66.7%; Pred. No. 0.35;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 2 SKNPHLOXNTIG 13
 23 SKAHLQKSTGC 34
 RESULT 6
 ID_CPS1_YEAST STANDARD; PRT; 530 AA.
 AC P10614;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome P450 51 (EC 1.14.14.-) (CYPL1) (P450-L1A1) (Sterol 14-alpha demethylase) (lanosterol 14-alpha demethylase) (P450-14DM).
 GN ERK11 OR CYP51 OR YHR007C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88111027; PubMed=3322742;
 RA Kalb V.F., Woods C.W., Turi T.G., Dey C.R., Sutter T.R., Loper J.C.;
 RT "Primary structure of the P450 lanosterol demethylase gene from
 RT Saccharomyces cerevisiae.";
 RL DNA 6:529-537 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88326319; PubMed=3046615;
 RA Ishida N., Aoyama Y., Hatanaka R., Oyama Y., Imajo S., Ishiguro M.,
 RA Oshida T., Nakazato H., Noguchi T., Maitra U.S., Mohan V.P.,
 RA Sprinson D.B., Yoshida Y.;
 RT "A single amino acid substitution converts cytochrome P450 (14DM) to
 RT an inactive form, cytochrome P450G1: complete primary structures
 RT deduced from cloned DNAs.";
 RL Biochem. Biophys. Res. Commun. 155:317-323 (1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Faveille A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
 RA Nham M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VII.";
 RL Science 265:2077-2082 (1994).
 RN [4]
 RP SEQUENCE OF 444-530 FROM N.A.
 RX MEDLINE=87106820; PubMed=3542713;
 RA Kalb V.F., Loper J.C., Dey C.R., Woods C.W., Sutter T.R.;
 RT "Isolation of a cytochrome P-450 structural gene from Saccharomyces
 RT cerevisiae.";
 RL Gene 45:237-245 (1986).
 CC -1- FUNCTION: CATALYZES C14-DEMETHYLATION OF LANOSTEROL WHICH IS
 CC CRITICAL FOR ERGOSTEROL BIOSYNTHESIS. IT TRANSFORMS LANOSTEROL
 CC INTO 4,4'-DIMETHYL CHOLESTA-8,14,24-TRIENE-3-BETA-OL.
 CC -1- PATHWAY: Ergosterol biosynthesis.
 CC -1- MISCELLANEOUS: IT IS THE MAIN TARGET FOR ANTIFUNGAL COMPOUNDS OF
 CC THE TRIAZOLE FAMILY LIKE KETOCONAZOLE WHICH INHIBITS BY
 CC COORDINATING THE IRON ATOM AT THE SIXTH LIGAND POSITION.

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M18109; AAA34379.1; -
 CC EMBL; M15663; AAA34837.1; -
 CC EMBL; M21483; AAA34546.1; -
 CC EMBL; M21484; AAA34547.1; -
 CC EMBL; U10555; AAB68433.1; -
 CC PIR; A25563; A25563.
 CC PIR; A27491; A27491.
 CC PIR; B31569; B31569.
 CC SGD; S0001049; ERG11.
 CC InterPro: IPR001128; Cytochrome_P450.
 CC Pfam: PF00067; P450.1.
 CC DR PRINTS; PR00385; P450.
 CC DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 CC KW Electron transport; Oxidoreductase; Monooxygenase; Membrane; Heme;
 CC Sterol biosynthesis.
 CC FT BINDING 470 HEME (BY SIMILARITY).
 CC FT CONFLICT 433 K -> N (IN REF. 2).
 CC SQ SEQUENCE 530 AA; 60720 MW; 646960BBA0E17979 CRC64;
 Query Match 51.9%; Score 41; DB 1; Length 530;
 Best Local Similarity 61.5%; Pred. No. 5.7;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 OY 2 SKNPHLOXNTIG 14
 181 SKNPHLMEKTTGT 193
 Db 181 SKNPHLMEKTTGT 193
 RESULT 7
 ID_Y364_RICPR STANDARD; PRT; 306 AA.
 AC Q92D63;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein RP364.
 GN RP364.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OC NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Anderson S.G.E., Zomrodipour A., Anderson J.O.,
 RA Scheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria.";
 RL Nature 396:133-140 (1998).
 CC -1- SIMILARITY: SOME, TO R. PROWAZEKII RP363.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ235271; CAA14823.1; -
 CC DR Hypothetical protein; Complete proteome.

SQ SEQUENCE 386 AA; 44346 MW; AAEE08EBBD43A0E7 CRC64;
 Query Match 50.6%; Score 40; DB 1; Length 386;
 Best Local Similarity 54.5%; Pred. No. 6.2;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 5 PHLOKNTIGTG 15
 :|||:|:|:|:
 Db 214 FHMKNVNEG 224

RESULT 8

AMYG YEAST STANDARD; PRT; 549 AA.
 AC P08019;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Glucosylase, intracellular sporulation-specific (EC 3.2.1.3) (glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).
 GN SGAL OR SGA OR YIL099W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=87194600; PubMed=3106330;
 RA Yamashita I., Nakamura M., Fukui S.;
 RT "Gene fusion is a possible mechanism underlying the evolution of
 RT STR1.";
 RL J. Bacteriol. 169:2142-2149 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentsch S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- SIMILARITY: BELONGS TO FAMILY 15 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z38125; CAA86282.1; -;
 DR EMBL; M16166; AAA35042.1; -;
 DR PIR; C26877; C26877.
 DR PIR; S48474; S48474.
 DR HSSP; P08017; IAYX.
 DR SGD; S0001361; SGAL.
 DR InterPro; IPR000165; GH 15.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PRINTS; PR00736; GLUTRILASE1.
 DR PROSITE; PS00820; GLUCOAMYLASE_1.
 DR KEGG; K00002; POLYSACCHARIDE DEGRADATION; Sporulation.
 KW Hydrolyase; Glycosidase; Polysaccharide degradation; Sporulation.
 FT BINDING 198 198 SUBSTRATE (BY SIMILARITY).
 FT ACT_SITE 261 261 CATALYTIC BASE (BY SIMILARITY).
 FT ACT_SITE 264 264 GENERAL ACID CATALYST (BY SIMILARITY).
 FT ACT_SITE 265 265 INTERACT WITH SUBSTRATE (BY SIMILARITY).
 FT CONFLICT 504 549 HVGDSLSPQFNKTCFMOGAGHILWYSFMDAYIOIRE
 FT VLOSL-> TWEOGN (IN REF. 1).
 SQ SEQUENCE 549 AA; 61463 MW; 6351b84f2cf4ab77 CRC64;

Query Match 50.6%; Score 40; DB 1; Length 549;
 Best Local Similarity 57.1%; Pred. No. 9.2;
 Matches 8; Conservative 3; Mismatches 1; Indels 2; Gaps 1;
 QY 4 NFHLK--NTIGTG 15
 :|||:|:|:|:
 Db 159 SFHLGRSNTLIGAG 172

RESULT 9

AMYH SAGDI STANDARD; PRT; 767 AA.
 AC P04065; O92314;
 DT 01-NOV-1986 (Rel. 03, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Glucosylase S1 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)
 DE (1,4-alpha-D-glucan glucohydrolase) (GAL).
 GN STR1 OR DEX2 OR MAL5.
 OS Saccharomyces diastaticus (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=41870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 60709;
 RX MEDLINE=85104778; PubMed=3918017;
 RA Yamashita I., Suzuki K., Fukui S.;
 RT "Nucleotide sequence of the extracellular glucosylase gene STR1 in
 RT the yeast Saccharomyces diastaticus.";
 RL J. Bacteriol. 161:567-573 (1985).
 RN [2]
 RP SEQUENCE OF 1-54 FROM N.A.
 RA Yamashita I., Suzuki K., Sakuzo F.;
 RT "Proteolytic processing of glucosylase in the yeast Saccharomyces
 RT cerevisiae.";
 RL Agric. Biol. Chem. 50:475-482 (1986).
 RN [3]
 RP SEQUENCE OF 1-64 FROM N.A.
 RA Shima H., Inui M., Akada R., Yamashita I.;
 RT "Upstream regions of the yeast glucosylase gene which are required
 RT for efficient transcription.";
 RL Agric. Biol. Chem. 53:749-755 (1989).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- SIMILARITY: BELONGS TO FAMILY 15 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02649; CAA26487.1; ALT_INIT.
 DR EMBL; D00428; BAAD0032.1; -;
 DR PIR; A21886; ALBYG.
 DR HSSP; P08017; IAYX.
 DR InterPro; IPR000165; GH 15.
 DR Pfam; PF00723; Glyco_hydro_15; 1.
 DR PROSITE; PS00820; GLUCOAMYLASE_1.
 DR KEGG; K00002; POLYSACCHARIDE DEGRADATION; Glycoprotein;
 KW Signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 767 GLUCOAMYLASE S1.
 FT DOMAIN 22 347 SER/THR-RICH.
 FT DOMAIN 348 691 H SUBUNIT.
 FT DOMAIN 692 767 Y SUBUNIT.
 FT BINDING 455 455 SUBSTRATE (BY SIMILARITY).
 FT ACT_SITE 518 518 CATALYTIC BASE (BY SIMILARITY).
 FT

FT ACT_SITE 521 521 GENERAL ACID CATALYST (BY SIMILARITY).
 FT ACT_SITE 522 522 INTERACT WITH SUBSTRATES (BY SIMILARITY).
 FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 645 645 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 650 650 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 720 720 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 741 741 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 767 AA; 82488 MW; A5F29E2427EDB593 CRC64;
 Query Match 50.6%; Score 40; DB 1; Length 767;
 Best Local Similarity 57.1%; Pred. No. 13;
 Matches 8; Conservative 3; Mismatches 1; Indels 2; Gaps 1;
 QY 4 NEHOK-NTIGTG 15
 Db 416 SFHLQRTNNTLGGG 429
 RESULT 10
 ID AMY1_SACDI STANDARD; PRT; 768 AA.
 AC P29760;
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Glucoamylase S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase)
 DE (1,4-alpha-D-glucan glucosyltransferase) (GAL1).
 GN STR2 OR DEX1.
 OS Saccharomyces diastaticus (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=41870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94288586; PubMed=8017901;
 RA Kim K., Bajszar G., Lee S.Y., Knudsen F., Mattoon J.R.;
 RT "Cloning of a new allelic variant of a Saccharomyces diastaticus
 RT glucosylase gene and its introduction into industrial yeasts.";
 RL Appl. Biochem. Biotechnol. 44:161-185(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91276266; PubMed=2055484;
 RA Lambrechts M.G., Pretorius I.S., Sollietti P., Marmur J.;
 RT "Primary structure and regulation of a glucosylase-encoding gene
 RT (STR2) in Saccharomyces diastaticus.";
 RL Gene 100:95-103(1991).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC -1- SIMILARITY: BELONGS TO FAMILY 15 OF GLYCOSYL HYDROLASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M60650; AAA35107.1; -;
 CC EMBL; M90490; AAA20560.1; -;
 CC PIR; S35895; S35895.
 CC HSSP; P08017; IAYX.
 CC InterPro; IPR000165; GH_15.
 CC Pfam; PF00723; Glyco_hydro_15; 1.
 CC PROSITE; PS00820; GLUCOAMYLASE; 1.

KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KW Signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 768
 FT DOMAIN 22 348
 FT DOMAIN 349 692
 FT DOMAIN 693 768
 FT BINDING 456 456
 FT ACT_SITE 519 519
 FT ACT_SITE 522 522
 FT ACT_SITE 523 523
 FT CARBOHYD 309 309
 FT CARBOHYD 323 323
 FT CARBOHYD 415 415
 FT CARBOHYD 424 424
 FT CARBOHYD 435 435
 FT CARBOHYD 514 514
 FT CARBOHYD 547 547
 FT CARBOHYD 646 646
 FT CARBOHYD 651 651
 FT CARBOHYD 721 721
 FT CARBOHYD 742 742
 FT CONFLICT 164 164
 FT CONFLICT 624 624
 SQ SEQUENCE 768 AA; 82586 MW; 3FAC172C128A0C6F CRC64;
 Query Match 50.6%; Score 40; DB 1; Length 768;
 Best Local Similarity 57.1%; Pred. No. 13;
 Matches 8; Conservative 3; Mismatches 1; Indels 2; Gaps 1;
 QY 4 NEHOK-NTIGTG 15
 Db 417 SFHLQRTNNTLGGG 430
 RESULT 11
 ID A29B_DROME STANDARD; PRT; 234 AA.
 AC O46157; Q9TWY3; Q9TW05; Q9TW06; Q9TW07; Q9U976; Q9U977; Q9U978;
 AC Q9U979; Q9V305;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Accessory gland protein Acp29Ab precursor.
 GN ACP29AB OR CG1797.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC STRAIN=Canton-S; TISSUE=Male accessory gland;
 RA MEDLINE=98135120; PubMed=9474779;
 RA Wolfner M.F., Harada H.A., Berttram M.J., Steelick T.J., Kraus K.W.,
 RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Tram U.K.;
 RT "New genes for male accessory gland proteins in Drosophila
 RT melanogaster.";
 RL Insect Biochem. Mol. Biol. 27:825-834(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Various strains;
 RX MEDLINE=99282496; PubMed=10353896;
 RA Aguade M.;
 RT "Positive selection drives the evolution of the Acp29Ab accessory
 RT gland protein in Drosophila.";
 RL Genetics 152:543-551(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Various strains;
 RX MEDLINE=2056153; PubMed=1102381;
 RA Begun D.J., Whitely P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;

"Molecular population genetics of male accessory gland proteins in *Drosophila*.";
 Genetics 156:1879-1888 (2000).
 [4]
 RC SEQUENCE FROM N.A.
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abail J.P., Agbayanti A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Borkakov S.,
 RA Borkova D., Botchan M.R., Bouck U., Brocktein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley D.A., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fodor C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.W., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskren D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 Science 287:2185-2195(2000).
 CC -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN
 CC MALE FEMALE FLIES.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLAND AND IN
 CC SEMINAL FLUID.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL, AJ240525; CAB53199.1;
 DR EMBL, AJ240526; CAB53200.1;
 DR EMBL, AJ240527; CAB53201.1;
 DR EMBL, AJ240528; CAB53202.1;
 DR EMBL, AJ240529; CAB53203.1;
 DR EMBL, AJ240530; CAB53204.1;
 DR EMBL, AJ240531; CAB53205.1;
 DR EMBL, AJ240532; CAB53206.1;
 DR EMBL, AJ240533; CAB53207.1;
 DR EMBL, AJ240534; CAB53208.1;
 DR EMBL, AJ240535; CAB53209.1;
 DR EMBL, AJ240536; CAB53210.1;
 DR EMBL, AJ240537; CAB53211.1;
 DR EMBL, AJ240538; CAB53212.1;
 DR EMBL, AJ240539; CAB53213.1;
 DR EMBL, AJ240540; CAB53214.1;
 DR EMBL, AJ240541; CAB53215.1;
 DR EMBL, AJ240542; CAB53216.1;
 DR EMBL, AJ240543; CAB53217.1;
 DR EMBL, AJ240544; CAB53218.1;
 DR EMBL, AJ240545; CAB53219.1;
 DR EMBL, AJ240546; CAB53220.1;
 DR EMBL, AJ240547; CAB53221.1;
 DR EMBL, AJ240548; CAB53222.1;
 DR EMBL, AJ240549; CAB53223.1;
 DR EMBL, AJ240550; CAB53224.1;
 DR EMBL, AJ240551; CAB53225.1;
 DR EMBL, AJ240552; CAB53259.1;
 DR EMBL, AJ240553; CAB53260.1;
 DR EMBL, AJ240554; CAB53261.1;
 DR EMBL, AJ240555; CAB53262.1;
 DR EMBL, AJ240556; CAB53263.1;
 DR EMBL, AJ240557; CAB53264.1;
 DR EMBL, AJ240558; CAB53265.1;
 DR EMBL, AJ240559; CAB53266.1;
 DR EMBL, AJ240560; CAB53267.1;
 DR EMBL, AJ240561; CAB53268.1;
 DR EMBL, AJ240562; CAB53269.1;
 DR EMBL, AJ240563; CAB53270.1;
 DR EMBL, AJ240564; CAB53271.1;
 DR EMBL, AJ240565; CAB53272.1;
 DR EMBL, AJ240566; CAB53273.1;
 DR EMBL, AJ240567; CAB53274.1;
 DR EMBL, AJ240568; CAB53275.1;
 DR EMBL, AJ240569; CAB53276.1;
 DR EMBL, AJ240570; CAB53277.1;
 DR EMBL, AJ240571; CAB53278.1;
 DR EMBL, AJ240572; CAB53279.1;
 DR EMBL, AJ240573; CAB53280.1;
 DR EMBL, AJ240574; CAB53281.1;
 DR EMBL, AJ240575; CAB53282.1;
 DR EMBL, AJ240576; CAB53283.1;
 DR EMBL, AJ240577; CAB53284.1;
 DR EMBL, AJ240578; CAB53285.1;
 DR EMBL, AJ240579; CAB53286.1;
 DR EMBL, AJ240580; CAB53287.1;
 DR EMBL, AJ240581; CAB53288.1;
 DR EMBL, AJ240582; CAB53289.1;
 DR EMBL, AJ240583; CAB53290.1;
 DR EMBL, AJ240584; CAB53291.1;
 DR EMBL, AJ240585; CAB53292.1;
 DR EMBL, AJ240586; CAB53293.1;
 DR EMBL, AJ240587; CAB53294.1;
 DR EMBL, AJ240588; CAB53295.1;
 DR EMBL, AJ240589; CAB53296.1;
 DR EMBL, AJ240590; CAB53297.1;
 DR EMBL, AJ240591; CAB53298.1;
 DR EMBL, AJ240592; CAB53299.1;
 DR EMBL, AJ240593; CAB53300.1;
 DR EMBL, AJ240594; CAB53301.1;
 DR EMBL, AJ240595; CAB53302.1;
 DR EMBL, AJ240596; CAB53303.1;
 DR EMBL, AJ240597; CAB53304.1;
 DR EMBL, AJ240598; CAB53305.1;
 DR EMBL, AJ240599; CAB53306.1;
 DR EMBL, AJ240600; CAB53307.1;
 DR EMBL, AJ240601; CAB53308.1;
 DR EMBL, AJ240602; CAB53309.1;
 DR EMBL, AJ240603; CAB53310.1;
 DR EMBL, AJ240604; CAB53311.1;
 DR EMBL, AJ240605; CAB53312.1;
 DR EMBL, AJ240606; CAB53313.1;
 DR EMBL, AJ240607; CAB53314.1;
 DR EMBL, AJ240608; CAB53315.1;
 DR EMBL, AJ240609; CAB53316.1;
 DR EMBL, AJ240610; CAB53317.1;
 DR EMBL, AJ240611; CAB53318.1;
 DR EMBL, AJ240612; CAB53319.1;
 DR EMBL, AJ240613; CAB53320.1;
 DR EMBL, AJ240614; CAB53321.1;
 DR EMBL, AJ240615; CAB53322.1;
 DR EMBL, AJ240616; CAB53323.1;
 DR EMBL, AJ240617; CAB53324.1;
 DR EMBL, AJ240618; CAB53325.1;
 DR EMBL, AJ240619; CAB53326.1;
 DR EMBL, AJ240620; CAB53327.1;
 DR EMBL, AJ240621; CAB53328.1;
 DR EMBL, AJ240622; CAB53329.1;
 DR EMBL, AJ240623; CAB53330.1;
 DR EMBL, AJ240624; CAB53331.1;
 DR EMBL, AJ240625; CAB53332.1;
 DR EMBL, AJ240626; CAB53333.1;
 DR EMBL, AJ240627; CAB53334.1;
 DR EMBL, AJ240628; CAB53335.1;
 DR EMBL, AJ240629; CAB53336.1;
 DR EMBL, AJ240630; CAB53337.1;
 DR EMBL, AJ240631; CAB53338.1;
 DR EMBL, AJ240632; CAB53339.1;
 DR EMBL, AJ240633; CAB53340.1;
 DR EMBL, AJ240634; CAB53341.1;
 DR EMBL, AJ240635; CAB53342.1;
 DR EMBL, AJ240636; CAB53343.1;
 DR EMBL, AJ240637; CAB53344.1;
 DR EMBL, AJ240638; CAB53345.1;
 DR EMBL, AJ240639; CAB53346.1;
 DR EMBL, AJ240640; CAB53347.1;
 DR EMBL, AJ240641; CAB53348.1;
 DR EMBL, AJ240642; CAB53349.1;
 DR EMBL, AJ240643; CAB53350.1;
 DR EMBL, AJ240644; CAB53351.1;
 DR EMBL, AJ240645; CAB53352.1;
 DR EMBL, AJ240646; CAB53353.1;
 DR EMBL, AJ240647; CAB53354.1;
 DR EMBL, AJ240648; CAB53355.1;
 DR EMBL, AJ240649; CAB53356.1;
 DR EMBL, AJ240650; CAB53357.1;
 DR EMBL, AJ240651; CAB53358.1;
 DR EMBL, AJ240652; CAB53359.1;
 DR EMBL, AJ240653; CAB53360.1;
 DR EMBL, AJ240654; CAB53361.1;
 DR EMBL, AJ240655; CAB53362.1;
 DR EMBL, AJ240656; CAB53363.1;
 DR EMBL, AJ240657; CAB53364.1;
 DR EMBL, AJ240658; CAB53365.1;
 DR EMBL, AJ240659; CAB53366.1;
 DR EMBL, AJ240660; CAB53367.1;
 DR EMBL, AJ240661; CAB53368.1;
 DR EMBL, AJ240662; CAB53369.1;
 DR EMBL, AJ240663; CAB53370.1;
 DR EMBL, AJ240664; CAB53371.1;
 DR EMBL, AJ240665; CAB53372.1;
 DR EMBL, AJ240666; CAB53373.1;
 DR EMBL, AJ240667; CAB53374.1;
 DR EMBL, AJ240668; CAB53375.1;
 DR EMBL, AJ240669; CAB53376.1;
 DR EMBL, AJ240670; CAB53377.1;
 DR EMBL, AJ240671; CAB53378.1;
 DR EMBL, AJ240672; CAB53379.1;
 DR EMBL, AJ240673; CAB53380.1;
 DR EMBL, AJ240674; CAB53381.1;
 DR EMBL, AJ240675; CAB53382.1;
 DR EMBL, AJ240676; CAB53383.1;
 DR EMBL, AJ240677; CAB53384.1;
 DR EMBL, AJ240678; CAB53385.1;
 DR EMBL, AJ240679; CAB53386.1;
 DR EMBL, AJ240680; CAB53387.1;
 DR EMBL, AJ240681; CAB53388.1;
 DR EMBL, AJ240682; CAB53389.1;
 DR EMBL, AJ240683; CAB53390.1;
 DR EMBL, AJ240684; CAB53391.1;
 DR EMBL, AJ240685; CAB53392.1;
 DR EMBL, AJ240686; CAB53393.1;
 DR EMBL, AJ240687; CAB53394.1;
 DR EMBL, AJ240688; CAB53395.1;
 DR EMBL, AJ240689; CAB53396.1;
 DR EMBL, AJ240690; CAB53397.1;
 DR EMBL, AJ240691; CAB53398.1;
 DR EMBL, AJ240692; CAB53399.1;
 DR EMBL, AJ240693; CAB53400.1;
 DR EMBL, AJ240694; CAB53401.1;
 DR EMBL, AJ240695; CAB53402.1;
 DR EMBL, AJ240696; CAB53403.1;
 DR EMBL, AJ240697; CAB53404.1;
 DR EMBL, AJ240698; CAB53405.1;
 DR EMBL, AJ240699; CAB53406.1;
 DR EMBL, AJ240700; CAB53407.1;
 DR EMBL, AJ240701; CAB53408.1;
 DR EMBL, AJ240702; CAB53409.1;
 DR EMBL, AJ240703; CAB53410.1;
 DR EMBL, AJ240704; CAB53411.1;
 DR EMBL, AJ240705; CAB53412.1;
 DR EMBL, AJ240706; CAB53413.1;
 DR EMBL, AJ240707; CAB53414.1;
 DR EMBL, AJ240708; CAB53415.1;
 DR EMBL, AJ240709; CAB53416.1;
 DR EMBL, AJ240710; CAB53417.1;
 DR EMBL, AJ240711; CAB53418.1;
 DR EMBL, AJ240712; CAB53419.1;
 DR EMBL, AJ240713; CAB53420.1;
 DR EMBL, AJ240714; CAB53421.1;
 DR EMBL, AJ240715; CAB53422.1;
 DR EMBL, AJ240716; CAB53423.1;
 DR EMBL, AJ240717; CAB53424.1;
 DR EMBL, AJ240718; CAB53425.1;
 DR EMBL, AJ240719; CAB53426.1;
 DR EMBL, AJ240720; CAB53427.1;
 DR EMBL, AJ240721; CAB53428.1;
 DR EMBL, AJ240722; CAB53429.1;
 DR EMBL, AJ240723; CAB53430.1;
 DR EMBL, AJ240724; CAB53431.1;
 DR EMBL, AJ240725; CAB53432.1;
 DR EMBL, AJ240726; CAB53433.1;
 DR EMBL, AJ240727; CAB53434.1;
 DR EMBL, AJ240728; CAB53435.1;
 DR EMBL, AJ240729; CAB53436.1;
 DR EMBL, AJ240730; CAB53437.1;
 DR EMBL, AJ240731; CAB53438.1;
 DR EMBL, AJ240732; CAB53439.1;
 DR EMBL, AJ240733; CAB53440.1;
 DR EMBL, AJ240734; CAB53441.1;
 DR EMBL, AJ240735; CAB53442.1;
 DR EMBL, AJ240736; CAB53443.1;
 DR EMBL, AJ240737; CAB53444.1;
 DR EMBL, AJ240738; CAB53445.1;
 DR EMBL, AJ240739; CAB53446.1;
 DR EMBL, AJ240740; CAB53447.1;
 DR EMBL, AJ240741; CAB53448.1;
 DR EMBL, AJ240742; CAB53449.1;
 DR EMBL, AJ240743; CAB53450.1;
 DR EMBL, AJ240744; CAB53451.1;
 DR EMBL, AJ240745; CAB53452.1;
 DR EMBL, AJ240746; CAB53453.1;
 DR EMBL, AJ240747; CAB53454.1;
 DR EMBL, AJ240748; CAB53455.1;
 DR EMBL, AJ240749; CAB53456.1;
 DR EMBL, AJ240750; CAB53457.1;
 DR EMBL, AJ240751; CAB53458.1;
 DR EMBL, AJ240752; CAB53459.1;
 DR EMBL, AJ240753; CAB53460.1;
 DR EMBL, AJ240754; CAB53461.1;
 DR EMBL, AJ240755; CAB53462.1;
 DR EMBL, AJ240756; CAB53463.1;
 DR EMBL, AJ240757; CAB53464.1;
 DR EMBL, AJ240758; CAB53465.1;
 DR EMBL, AJ240759; CAB53466.1;
 DR EMBL, AJ240760; CAB53467.1;
 DR EMBL, AJ240761; CAB53468.1;
 DR EMBL, AJ240762; CAB53469.1;
 DR EMBL, AJ240763; CAB53470.1;
 DR EMBL, AJ240764; CAB53471.1;
 DR EMBL, AJ240765; CAB53472.1;
 DR EMBL, AJ240766; CAB53473.1;
 DR EMBL, AJ240767; CAB53474.1;
 DR EMBL, AJ240768; CAB53475.1;
 DR EMBL, AJ240769; CAB53476.1;
 DR EMBL, AJ240770; CAB53477.1;
 DR EMBL, AJ240771; CAB53478.1;
 DR EMBL, AJ240772; CAB53479.1;
 DR EMBL, AJ240773; CAB53480.1;
 DR EMBL, AJ240774; CAB53481.1;
 DR EMBL, AJ240775; CAB53482.1;
 DR EMBL, AJ240776; CAB53483.1;
 DR EMBL, AJ240777; CAB53484.1;
 DR EMBL, AJ240778; CAB53485.1;
 DR EMBL, AJ240779; CAB53486.1;
 DR EMBL, AJ240780; CAB53487.1;
 DR EMBL, AJ240781; CAB53488.1;
 DR EMBL, AJ240782; CAB53489.1;
 DR EMBL, AJ240783; CAB53490.1;
 DR EMBL, AJ240784; CAB53491.1;
 DR EMBL, AJ240785; CAB53492.1;
 DR EMBL, AJ240786; CAB53493.1;
 DR EMBL, AJ240787; CAB53494.1;
 DR EMBL, AJ240788; CAB53495.1;
 DR EMBL, AJ240789; CAB53496.1;
 DR EMBL, AJ240790; CAB53497.1;
 DR EMBL, AJ240791; CAB53498.1;
 DR EMBL, AJ240792; CAB53499.1;
 DR EMBL, AJ240793; CAB53500.1;
 DR EMBL, AJ240794; CAB53501.1;
 DR EMBL, AJ240795; CAB53502.1;
 DR EMBL, AJ240796; CAB53503.1;
 DR EMBL, AJ240797; CAB53504.1;
 DR EMBL, AJ240798; CAB53505.1;
 DR EMBL, AJ240799; CAB53506.1;
 DR EMBL, AJ240800; CAB53507.1;
 DR EMBL, AJ240801; CAB53508.1;
 DR EMBL, AJ240802; CAB53509.1;
 DR EMBL, AJ240803; CAB53510.1;
 DR EMBL, AJ240804; CAB53511.1;
 DR EMBL, AJ240805; CAB53512.1;
 DR EMBL, AJ240806; CAB53513.1;
 DR EMBL, AJ240807; CAB53514.1;
 DR EMBL, AJ240808; CAB53515.1;
 DR EMBL, AJ240809; CAB53516.1;
 DR EMBL, AJ240810; CAB53517.1;
 DR EMBL, AJ240811; CAB53518.1;
 DR EMBL, AJ240812; CAB53519.1;
 DR EMBL, AJ240813; CAB53520.1;
 DR EMBL, AJ240814; CAB53521.1;
 DR EMBL, AJ240815; CAB53522.1;
 DR EMBL, AJ240816; CAB53523.1;
 DR EMBL, AJ240817; CAB53524.1;
 DR EMBL, AJ240818; CAB53525.1;
 DR EMBL, AJ240819; CAB53526.1;
 DR EMBL, AJ240820; CAB53527.1;
 DR EMBL, AJ240821; CAB53528.1;
 DR EMBL, AJ240822; CAB53529.1;
 DR EMBL, AJ240823; CAB53530.1;
 DR EMBL, AJ240824; CAB53531.1;
 DR EMBL, AJ240825; CAB53532.1;
 DR EMBL, AJ240826; CAB53533.1;
 DR EMBL, AJ240827; CAB53534.1;
 DR EMBL, AJ240828; CAB53535.1;
 DR EMBL, AJ240829; CAB53536.1;
 DR EMBL, AJ240830; CAB53537.1;
 DR EMBL, AJ240831; CAB53538.1;
 DR EMBL, AJ240832; CAB53539.1;
 DR EMBL, AJ240833; CAB53540.1;
 DR EMBL, AJ240834; CAB53541.1;
 DR EMBL, AJ240835; CAB53542.1;
 DR EMBL, AJ240836; CAB53543.1;
 DR EMBL, AJ240837; CAB53544.1;
 DR EMBL, AJ240838; CAB53545.1;
 DR EMBL, AJ240839; CAB53546.1;
 DR EMBL, AJ240840; CAB53547.1;
 DR EMBL, AJ240841; CAB53548.1;
 DR EMBL, AJ240842; CAB53549.1;
 DR EMBL, AJ240843; CAB53550.1;
 DR EMBL, AJ240844; CAB53551.1;
 DR EMBL, AJ240845; CAB53552.1;
 DR EMBL, AJ240846; CAB53553.1;
 DR EMBL, AJ240847; CAB53554.1;
 DR EMBL, AJ240848; CAB53555.1;
 DR EMBL, AJ240849; CAB53556.1;
 DR EMBL, AJ240850; CAB53557.1;
 DR EMBL, AJ240851; CAB53558.1;
 DR EMBL, AJ240852; CAB53559.1;
 DR EMBL, AJ240853; CAB53560.1;
 DR EMBL, AJ240854; CAB53561.1;
 DR EMBL, AJ240855; CAB53562.1;
 DR EMBL, AJ240856; CAB53563.1;
 DR EMBL, AJ240857; CAB53564.1;
 DR EMBL, AJ240858; CAB53565.1;
 DR EMBL, AJ240859; CAB53566.1;
 DR EMBL, AJ240860; CAB53567.1;
 DR EMBL, AJ240861; CAB53568.1;
 DR EMBL, AJ240862; CAB53569.1;
 DR EMBL, AJ240863; CAB53570.1;
 DR EMBL, AJ240864; CAB53571.1;
 DR EMBL, AJ240865; CAB53572.1;
 DR EMBL, AJ240866; CAB53573.1;
 DR EMBL, AJ240867; CAB53574.1;
 DR EMBL, AJ240868; CAB53575.1;
 DR EMBL, AJ240869; CAB53576.1;
 DR EMBL, AJ240870; CAB53577.1;
 DR EMBL, AJ240871; CAB53578.1;
 DR EMBL, AJ240872; CAB53579.1;
 DR EMBL, AJ240873; CAB53580.1;
 DR EMBL, AJ240874; CAB53581.1;
 DR EMBL, AJ240875; CAB53582.1;
 DR EMBL, AJ240876; CAB53583.1;
 DR EMBL, AJ240877; CAB53584.1;
 DR EMBL, AJ240878; CAB53585.1;
 DR EMBL, AJ240879; CAB53586.1;
 DR EMBL, AJ240880; CAB53587.1;
 DR EMBL, AJ240881; CAB53588.1;
 DR EMBL, AJ240882; CAB53589.1;
 DR EMBL, AJ240883; CAB53590.1;
 DR EMBL, AJ240884; CAB53591.1;
 DR EMBL, AJ240885; CAB53592.1;
 DR EMBL, AJ240886; CAB53593.1;
 DR EMBL, AJ240887; CAB53594.1;
 DR EMBL, AJ240888; CAB53595.1;
 DR EMBL, AJ240889; CAB53596.1;
 DR EMBL, AJ240890; CAB53597.1;
 DR EMBL, AJ240891; CAB53598.1;
 DR EMBL, AJ240892; CAB53599.1;
 DR EMBL, AJ240893; CAB53600.1;
 DR EMBL, AJ240894; CAB53601.1;
 DR EMBL, AJ240895; CAB53602.1;
 DR EMBL, AJ240896; CAB53603.1;
 DR EMBL, AJ240897; CAB53604.1;
 DR EMBL, AJ240898; CAB53605.1;
 DR EMBL, AJ240899; CAB53606.1;
 DR EMBL, AJ240900; CAB53607.1;
 DR EMBL, AJ240901; CAB53608.1;
 DR EMBL, AJ240902; CAB53609.1;
 DR EMBL, AJ240903; CAB53610.1;
 DR EMBL, AJ240904; CAB53611.1;
 DR EMBL, AJ240905; CAB53612.1;
 DR EMBL, AJ240906; CAB53613.1;
 DR EMBL, AJ240907; CAB53614.1;
 DR EMBL, AJ240908; CAB53615.1;
 DR EMBL, AJ240909; CAB53616.1;
 DR EMBL, AJ240910; CAB53617.1;
 DR EMBL, AJ240911; CAB53618.1;
 DR EMBL, AJ240912; CAB53619.1;
 DR EMBL, AJ240913; CAB53620.1;
 DR EMBL, AJ240914; CAB53621.1;
 DR EMBL, AJ240915; CAB53622.1;
 DR EMBL, AJ240916; CAB53623.1;
 DR EMBL, AJ240917; CAB53624.1;
 DR EMBL, AJ240918; CAB53625.1;
 DR EMBL, AJ240919; CAB53626.1;
 DR EMBL, AJ240920; CAB53627.1;
 DR EMBL, AJ240921; CAB53628.1;
 DR EMBL, AJ240922; CAB53629.1;
 DR EMBL, AJ240923; CAB53630.1;
 DR EMBL, AJ240924; CAB53631.1;
 DR EMBL, AJ240925; CAB53632.1;
 DR EMBL, AJ240926; CAB53633.1;
 DR EMBL, AJ240927; CAB53634.1;
 DR EMBL, AJ240928; CAB53635.1;
 DR EMBL, AJ240929; CAB53636.1;
 DR EMBL, AJ240930; CAB53637.1;
 DR EMBL, AJ240931; CAB53638.1;
 DR EMBL, AJ240932; CAB53639.1;
 DR EMBL, AJ240933; CAB53640.1;
 DR EMBL, AJ240934; CAB53641.1;
 DR EMBL, AJ240935; CAB53642.1;
 DR EMBL, AJ240936; CAB53643.1;
 DR EMBL, AJ24

FT MA5, MA21, MA45, MA52, MA67, ZIM29,
ZIM30, ZIM42 AND ZIM56).
E -> D (IN STRAIN MAY7).
FT VARIANT 214 214
SQ SEQUENCE 234 AA; 27173 MW; 8954CD3215480F3E CRC64;

Query Match 49.4%; Score 39; DB 1; Length 234;
Best Local Similarity 46.2%; Pred. No. 5.5;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 SKNFHLQKNTGT 14
DB 119 SNFHLTKMLQGT 131

RESULT 12

QYEA_STRPN STANDARD; PRT; 342 AA.

AC Q97015; 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-adenosylmethionine:RNA ribosyltransferase-isomerase (EC 5.-.-.-)
GN (Quenosine biosynthesis protein quea).
OS QYEA OR SP1416.
OC Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIGR4;
RX MEDLINE=21357209; PubMed=1163916;
RA Tettein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heideberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Swinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayar L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Anguioni S., Dickinson T., Hickey E.K.,
RA Holt I.E., Lotius B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RU Science 293:498-506(2001).
CC -1- FUNCTION: Synthesizes oq from preqi in a single S-
CC adenosylmethionine-requiring step. The ribosyl moiety of Adomet is
CC transferred and isomerized to the epoxycyclopentane residue of oq
CC (By similarity).
CC -1- PATHWAY: Quenosine biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE QYEA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE007438; AAK75514.1; -.
CC TIGR; SP1416; -.
CC InterPro: IPR003699; Quenosine synth.
CC Pfam: PF02547; Quenosine synth. 1.
CC Trifam: TRIGR00113; quea.1.
CC DR TRIFAMs; TRIGR00113; quea.1.
CC KW Quenosine biosynthesis; transferase; isomerase; Complete proteome.
SQ SEQUENCE 342 AA; 38379 MW; E81265A1C36D59 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 342;
Best Local Similarity 63.6%; Pred. No. 8.4;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SKNFHLQKNTI 12
QY 119 SNFHLQKNTI 12

DB 295 STNFHLPKSTL 305

RESULT 13

QYEA_STRPY STANDARD; PRT; 342 AA.

AC Q99249; 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-adenosylmethionine:RNA ribosyltransferase-isomerase (EC 5.-.-.-)
GN (Quenosine biosynthesis protein quea).
OS QYEA OR SPY1400.
OC Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti V.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Szate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RU Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -1- FUNCTION: Synthesizes oq from preqi in a single S-
CC adenosylmethionine-requiring step. The ribosyl moiety of Adomet is
CC transferred and isomerized to the epoxycyclopentane residue of oq
CC (By similarity).
CC -1- PATHWAY: Quenosine biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE QYEA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006577; AAK34216.1; -.
CC InterPro: IPR003699; Quenosine synth.
CC Pfam: PF02547; Quenosine synth. 1.
CC DR TRIFAMs; TRIGR00113; quea.1.
CC KW Quenosine biosynthesis; transferase; isomerase; Complete proteome.
SQ SEQUENCE 342 AA; 38351 MW; E46CD9203E3FDBF2 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 342;
Best Local Similarity 63.6%; Pred. No. 8.4;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SKNFHLQKNTI 12
DB 295 STNFHLPKSTL 305

RESULT 14

PEPG_LACDL STANDARD; PRT; 437 AA.

AC P94869; 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Aminoacylase G (EC 3.4.22.-).
GN PEPG.
OS Lactobacillus delbrueckii (subsp. lactis).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=29397;
RN [1]


```

RP SEQUENCE FROM N.A.
RC STRAIN=DSM 7290;
RX MEDLINE=97195798; PubMed=9043129;
RA Klein J.R., Schick J., Henrich B., Plapp R.;
RT "necrobacillus delbrueckii subsp. lactis DSM7290 pepg gene encodes a
RL novel cysteine aminopeptidase."
CC Microbiology 143:527-537(1997).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z71782; CA96465.1; -
CC HSSP: Q13867; 2CB5.
CC DR MEROPS; C01.089; -
CC DR InterPro: IPR004134; Pept_C1 like.
CC DR InterPro: IPR000668; Peptidase C1.
CC DR InterPro: IPR000169; SHprot_acsite.
CC DR Pfam: PF03051; Pept_C1-like; 1.
CC DR PRINTS; PR00705; PAPAIN.
CC DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
CC DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
CC DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.
CC DR Hydrolase; Thiol protease; Aminopeptidase.
CC FT ACT_SITE 70 70 BY SIMILARITY.
CC FT ACT_SITE 361 361 BY SIMILARITY.
CC FT ACT_SITE 382 382 BY SIMILARITY.
CC SQ SEQUENCE 437 AA; 49705 MW; 2099C62E57181FC5 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 437;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 KNFHLQKN 10
DB 234 KNVHLEKN 241

RESULT 15
FLIV_SALMU STANDARD; PRT; 173 AA.
AC P37588;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flagellar biosynthetic protein fliv.
GN FLIV.
OS Salmonella muenchen.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=596;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 8388;
RX MEDLINE=93231524; PubMed=8472952;
RA Doll L., Frankel G.;
RT "Cloning and sequencing of two new flv genes, the products of which
RT are essential for Salmonella flagellar biosynthesis."
RL Gene 126:119-121(1993).
CC [2]
CC CHARACTERIZATION.
CC MEDLINE=94075958; PubMed=8254311;
CC Doll L., Frankel G.;
CC "fliv and fliv: two flagellar genes essential for biosynthesis of
CC Salmonella and Escherichia coli flagella."
CC J. Gen. Microbiol. 139:2415-2422(1993).
CC -1- FUNCTION: Required for the secretion of flagellin and expression

```

```

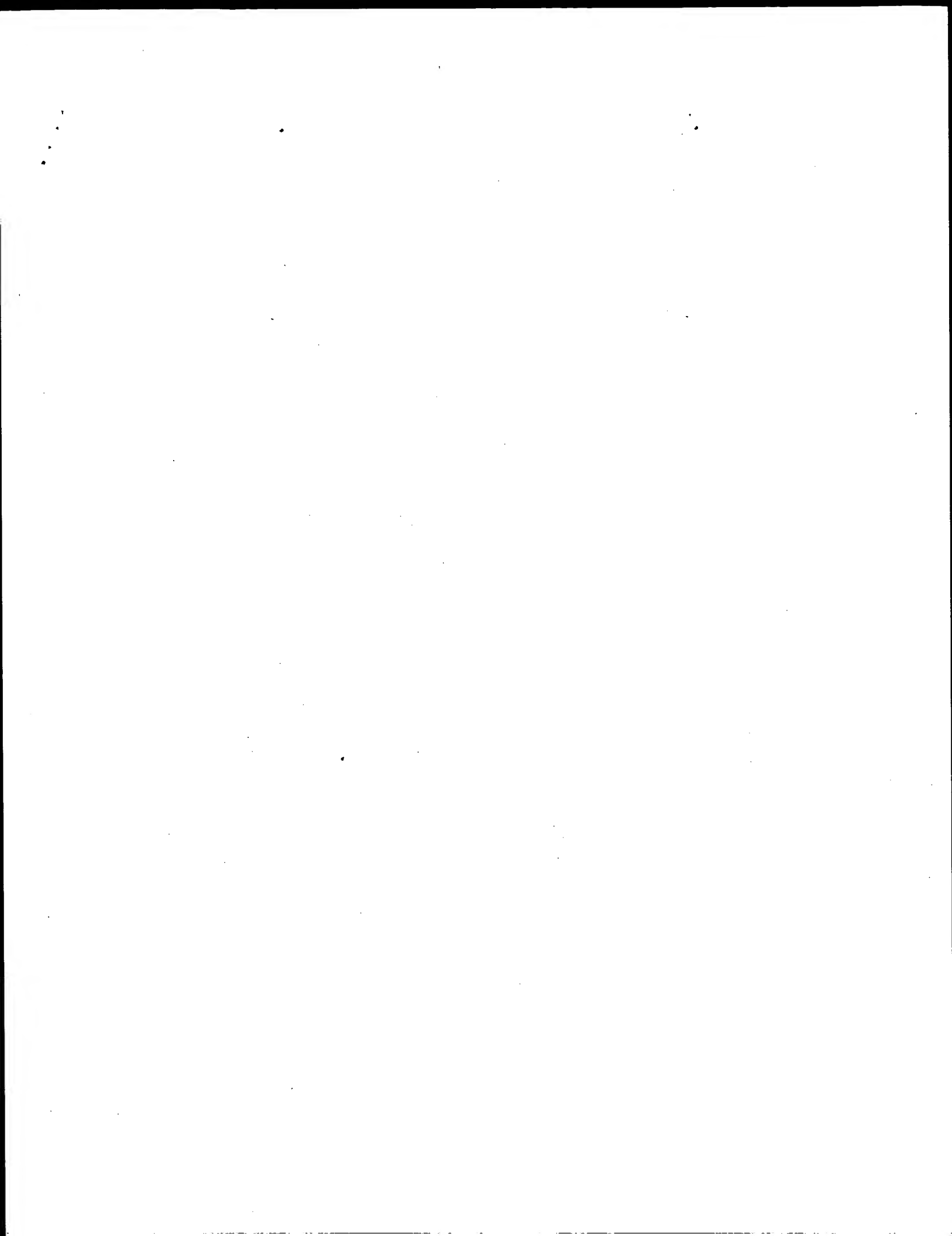
CC of motility.
CC -1- SIMILARITY: BELONGS TO THE FLIV FAMILY. CORRESPONDS TO THE C-
CC TERMINAL SECTION
CC -1- CAUTION: Unlike the other Salmonellae, the ortholog of the fliv
CC protein may be encoded by two CDS in S.muenchen. It cannot be
CC ruled out that sequencing errors produced two CDS instead of one.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L06521; AAA27107.1; -
CC DR PIR; JN0515; JN0515.
CC DR Flagella.
CC KW
CC SQ SEQUENCE 173 AA; 19982 MW; 84DFAB36D1FF32C8 CRC64;

Query Match 48.1%; Score 38; DB 1; Length 173;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SKNFHLQKNITGTV 15
DB 150 AKNFHRIETVATG 163

Search completed: April 20, 2003, 13:07:44
Job time : 3.92105 secs

```



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds

(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524d-131

Sequence: 1 SRAEVSYPHNGAKF 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR 73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	514	2	S48730
2	76	100.0	514	2	UC2498
3	63	82.9	514	2	UC7100
4	61	80.3	507	2	UC7366
5	45	59.2	155	2	E97507
6	45	59.2	155	2	A12725
7	43	56.6	419	1	J00346
8	43	56.6	431	2	T46187
9	43	56.6	456	2	E69391
10	42	55.3	131	2	S42733
11	42	55.3	385	2	S53838
12	42	55.3	386	2	S58749
13	41	53.9	198	2	C90417
14	41	53.9	316	2	G72077
15	41	53.9	316	2	G86545
16	41	53.9	1303	1	S27396
17	40	52.6	256	2	S34245
18	40	52.6	374	2	T11323
19	40	52.6	374	2	T11323
20	40	52.6	578	2	S68155
21	39	51.3	382	2	B64012
22	39	51.3	386	1	S12352
23	39	51.3	576	2	AC2195
24	39	51.3	4930	2	E69679
25	38	50.0	89	2	C89447
26	38	50.0	97	2	H69895
27	38	50.0	343	2	S62704
28	38	50.0	348	2	A70311
29	38	50.0	385	1	CBBY

ALIGNMENTS

RESULT 1

S48730

Cry j II protein - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999

C/Accession: S48730

R/Namba, M.; Kurose, M.; Torioka, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; FERS Lett. 353, 124-128, 1994

A/Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar

A/Reference number: S48730; PMID:95010777; PMID:7926035

A/Accession: S48730

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-514 <KMW>

A/Cross-References: GB:ID37765; NID:9577695; PIDN:BA07021.1; PID:dl007598; PID:9577694

Query Match 100.0%; Score 76; DB 2; Length 514;

Best Local Similarity 100.0%; Pred. No. 8.2e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAEVSYPHNGAKF 15

DB 290 SRAEVSYPHNGAKF 304

RESULT 2

second major allergen Cry j II precursor - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000

C/Accession: JC2346; PC2346; A60147

R/Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K. Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994

A/Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese cedar

A/Reference number: JC2498; PMID:94271186; PMID:8002972

A/Accession: JC2498

A/Molecule type: mRNA

A/Residues: 1-514 <KMW>

A/Cross-References: DBJ:ID29772; NID:9506857; PIDN:BA06172.1; PID:9506858

A/Accession: FC2346

A/Molecule type: protein

A/Residues: 52-61 <K02>

R/Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matunasi, T. Allergy 45, 309-312, 1990

A/Title: Identification of the second major allergen of Japanese cedar pollen.

A/Reference number: A60147; PMID:9034298; PMID:2382797

A/Accession: A60147

A/Molecule type: protein

A/Residues: 55-64 <SAK>

C/Keywords: glycoprotein; pollen

F:1-54/Domain: signal sequence

F:55-460/Product: second major allergen Cry j #status predicted <SIG>

F:55-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 76; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 8,2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAEVSYYHNGAKF 15
|||||
Db 290 SRAEVSYYHNGAKF 304

RESULT 3

JC7100 polygalacturonase Cha o 2 - Japanese cypress

C:Species: Chamaecyparis obtusa (Japanese cypress)

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000

C:Accession: JC7100; PC7026

R:Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.

Biochem. Biophys. Res. Commun. 263, 166-171, 1999

A:Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all

A:Reference number: JC7100; PMID:99417540; PMID:10486272

A:Accession: JC7100

A:Molecule type: mRNA

A:Residues: 1-514 <MOR>

A:Accession: PC7026

A:Molecule type: protein

A:Residues: 51-62 <MO2>

Query Match 82.9%; Score 63; DB 2; Length 514;

Best Local Similarity 86.7%; Pred. No. 0.0019;

Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 4

JC7366 Jun a 2 protein - mountain cedar

C:Species: Juniperus ashei (mountain cedar)

C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000

C:Accession: JC7366; PC7093

R:Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.

Biochem. Biophys. Res. Commun. 275, 195-202, 2000

A:Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all

A:Reference number: JC7366

A:Accession: JC7366

A:Molecule type: mRNA

A:Residues: 1-507 <YOK>

A:Cross-references: GB:AJ404653

A:Accession: PC7093

A:Molecule type: protein

A:Residues: 55-63 <YOK>

C:Comment: This protein, a second major allergen of mountain cedar pollen, which is inv

C:Keywords: glycoprotein; pollen

Query Match 80.3%; Score 61; DB 2; Length 507;

Best Local Similarity 73.3%; Pred. No. 0.0044;

Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

RESULT 5

E97507

hypothetical protein AGR_C_2237 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: E97507

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldma

A.; Liu, F.; McLiam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, E

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: E97507

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK87014.1; PID:G15156260; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C_2237

A:Map position: circular chromosome

Query Match 59.2%; Score 45; DB 2; Length 155;

Best Local Similarity 61.5%; Pred. No. 0.98;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 6

A12725

conserved hypothetical protein Atu1212 [imported] - Agrobacterium tumefaciens (strain

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: A12725

R:Wood, D.W.; Secubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCle

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: A12725

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 <KUR>

A:Cross-references: GB:AE008688; PIDN:AL42223.1; PID:G17739617; GSPDB:GN00186

C:Genetics:

A:Gene: Atu1212

A:Map position: circular chromosome

Query Match 59.2%; Score 45; DB 2; Length 155;

Best Local Similarity 61.5%; Pred. No. 0.98;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 7

JQ0346

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Rhodospseudomonas viridi

C:Species: Rhodospseudomonas viridis

C:Date: 07-Sep-1990 #sequence_revision 02-Aug-1994 #text_change 03-Jun-2002

C:Accession: JQ0346

R:Verbiest, J.; Lang, F.; Gabellini, N.; Oesterhelit, D.

Mol. Gen. Genet. 219, 445-452, 1989

A:Title: Cloning and sequencing of the fbcF, B and C genes encoding the cytochrome b

A:Reference number: JQ0345; PMID:90158506; PMID:2560136

A:Accession: JQ0346

A:Molecule type: DNA

A:Residues: 1-419 <VER>

A:Experimental source: strain 133

C:Comment: This protein is one of the three subunits of the ubiquinol-cytochrome C2 c

lectron-transport system that catalyzes the synthesis of ATP.

A:Gene: fbcb
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; oxidoreductase
 C:Accession: E69391
 F/24-356/Domain: cytochrome b homology <CB6>
 F/24-224/Domain: cytochrome b6 homology <CB6>
 F/49-65/Domain: transmembrane #status predicted <TM1>
 F/94-112/Domain: transmembrane #status predicted <TM2>
 F/132-148/Domain: transmembrane #status predicted <TM3>
 F/133-215/Domain: transmembrane #status predicted <TM4>
 F/238-356/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
 F/246-562/Domain: transmembrane #status predicted <TM5>
 F/305-321/Domain: transmembrane #status predicted <TM6>
 F/340-360/Domain: transmembrane #status predicted <TM7>
 F/370-386/Domain: transmembrane #status predicted <TM8>
 F/96-157/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F/110-211/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 56.6%; Score 43; DB 1; Length 419;
 Best Local Similarity 54.5%; Pred. No. 6.8;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 VSYVHNGAKF 15
 Db 92 IRYIHNGASF 102

RESULT 8
 T46187
 polygalacturonase (EC 3.2.1.15) precursor [similarity] - Arabidopsis thaliana
 N/Alternate names: protein T8H10.110
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 18-Aug-2000
 C:Accession: T46187; T50674
 R/Gene: V.; Rechmann, S.; Borkova, D.; Ansoyge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23014
 A:Accession: T46187
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-431 <BEN>
 A:Cross-references: EMBL:AL133248; PIDN:CAB6108.1
 A:Experimental source: cultivar Columbia; BAC clone T8H10
 R/Jenkins, E.S.; Roberts, J.A.
 submitted to the EMBL Data Library, December 1997
 A:Description: Dendrose-related expression of an Arabidopsis thaliana gene encoding a
 A:Reference number: Z25172
 A:Accession: T50674
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-431 <JEN>
 A:Cross-references: EMBL:AF037367; PIDN:AAC98923.1
 A:Experimental source: cultivar Landsberg erecta
 C:Genetics:
 A:Gene: T8H10.110
 A:Map position: 3
 A:Introns: 85/3; 129/3; 186/3; 193/3; 263/1; 290/2; 326/3; 364/3
 C:Superfamily: polygalacturonase
 C:Keywords: glycosidase; hydrolase
 F/1-27/Domain: signal sequence #status predicted <SIG>
 F/28-431/Product: polygalacturonase #status predicted <MAT>

Query Match 56.6%; Score 43; DB 2; Length 431;
 Best Local Similarity 60.0%; Pred. No. 7.1;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRRAVSIVHNGAKF 15
 Db 299 SKAYVGINVDGAKF 313

RESULT 9
 E69391
 hypothetical protein AF1134 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: E69391
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes,
 Smith, H.O.; Moore, C.R.; Venter, J.C.
 A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: E69391
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-456 <KLE>
 A:Cross-references: GB:AE001026; GB:AE000782; NID:G2689349; PIDN:AB90123.1; PID:G264

Query Match 56.6%; Score 43; DB 2; Length 456;
 Best Local Similarity 72.7%; Pred. No. 7.5;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 VSYVHNGAKF 15
 Db 178 VSYVLDGAKF 188

RESULT 10
 S42733
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - yeast (Williopsis mrakii
 C:Species: mitochondrion Williopsis mrakii
 C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 03-Jun-2002
 C:Accession: S42733
 R/Diressi, R.; Sor, F.; Nosek, J.; Fukuhara, H.
 Yeast 10, 391-398, 1994
 A>Title: Genes of the linear mitochondrial DNA of Williopsis mrakii: coding sequences

A:Reference number: S42733; MUID:94287716; PMID:8017108
 A:Accession: S42733
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-131 <DR>
 A:Cross-references: EMBL:X66594; NID:G296849; PIDN:CAA47157.1; PID:G296850
 A:Experimental source: strain CBS1707
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1992
 C:Genetics:
 A:Gene: CYTB
 A:Genome: mitochondrion
 A:Genetic code: SGC2
 A:Introns: 67/3
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin
 C:Keywords: chromoprotein; electron transfer; heme; iron; mitochondrion; oxidative pho
 F/10-131/Domain: cytochrome b6 homology (fragment) <CB6>
 F/10-131/Domain: cytochrome b homology (fragment) <CB6>
 F/35-51/Domain: transmembrane #status predicted <TM1>
 F/80-98/Domain: transmembrane #status predicted <TM2>

Query Match 55.3%; Score 42; DB 2; Length 131;
 Best Local Similarity 54.5%; Pred. No. 2.9;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 VSYVHNGAKF 15
 Db 78 IRYIHNGASF 88

RESULT 11
 S53838
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Acanthamoeba castellanii
 C:Species: mitochondrion Acanthamoeba castellanii
 C>Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 03-Jun-2002
 C:Accession: S53838
 R/Burger, G.; Plante, I.; Loneragan, K.M.; Gray, M.W.
 J. Mol. Biol. 245, 522-537, 1995

A>Title: The mitochondrial DNA of the amoeboid protozoan, *Acanthamoeba castellanii*: comp
 A:Reference number: S53825; MUID:95147275; PMID:7844823
 A:Accession: S53838
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-385 <BUR>
 A:Cross-references: GB:U12386; NID:9562028; PIDN:AA01830.1; PID:9562042
 A:Experimental source: strain Neff; ATCC 30010
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
 C:Genetics:

A:Gene: cob
 A:Genome: mitochondrion
 A:Genetic code: SGC6
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:10-341/Domain: cytochrome b6 homology <CB6>
 F:10-210/Domain: cytochrome b6 homology <CB6>
 F:35-51/Domain: transmembrane #status predicted <TM1>
 F:80-98/Domain: transmembrane #status predicted <TM2>
 F:118-134/Domain: transmembrane #status predicted <TM3>
 F:179-201/Domain: transmembrane #status predicted <TM4>
 F:223-341/Domain: plastocyanin reductase 17K protein homology <17K>
 F:231-247/Domain: transmembrane #status predicted <TM5>
 F:299-306/Domain: transmembrane #status predicted <TM6>
 F:325-345/Domain: transmembrane #status predicted <TM7>
 F:355-371/Domain: transmembrane #status predicted <TM8>
 F:82-183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:96-197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 55.3%; Score 42; DB 2; Length 385;
 Best Local Similarity 54.5%; Pred. No. 9.5;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 VSVYVHNGAKF 15
 : : : : :
 DB 78 IRYIHANGASF 88

RESULT 12
 S58749
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - yeast (*Hansenula wingei*) m
 C:Species: mitochondrion *Hansenula wingei*
 C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 03-Jun-2002
 C:Accession: S58749
 R:Seikito, T.; Okamoto, K.; Kitano, H.; Yoshida, K.
 Curr. Genet. 28, 39-53, 1995
 A>Title: The complete mitochondrial DNA sequence of *Hansenula wingei* reveals new charact
 A:Reference number: S58740; MUID:96022424; PMID:8536312
 A:Accession: S58749
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-386 <SEK>
 A:Cross-references: EMBL:D31785
 C:Genetics:

A:Genome: mitochondrion
 A:Genetic code: SGC3
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin
 C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
 F:10-340/Domain: cytochrome b6 homology <CB6>
 F:10-210/Domain: cytochrome b6 homology <CB6>
 F:223-340/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
 F:82-183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:96-197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 55.3%; Score 42; DB 2; Length 386;
 Best Local Similarity 54.5%; Pred. No. 9.5;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 VSVYVHNGAKF 15
 : : : : :
 DB 78 IRYIHANGASF 88

RESULT 13
 C90417
 endonuclease V (nfi) [imported] - *Sulfolobus solfataricus*
 C:Species: *Sulfolobus solfataricus*
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
 C:Accession: C90417
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awaiz, M.J.; Ch
 Jung, I.; Ueffing, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
 arrett, R.A.; Regan, W.A.; Sersen, C.W.; Van der Oost, J.
 submitted to Genbank, April 2001
 A:Description: *Sulfolobus solfataricus* complete genome.
 A:Reference number: A99139
 A:Accession: C90417
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-198 <RUR>
 A:Cross-references: GB:AE00641; NID:G13815757; PIDN:AAK42594.1; GSPDB:GN00155
 C:Genetics:

A:Superfamily: conserved hypothetical protein AF0129
 Query Match 53.9%; Score 41; DB 2; Length 198;
 Best Local Similarity 42.9%; Pred. No. 6.9;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAEVSYVHNGAK 14
 : : : : :
 DB 139 NESITVYVNGEK 152

RESULT 14
 G72077
 ct007 hypothetical protein - *Chlamydomonas reinhardtii* (strain CML029)
 C:Species: *Chlamydomonas reinhardtii*
 C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
 C:Accession: G72077
 R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
 Nature Genet. 21, 385-389, 1999
 A>Title: Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.
 A:Reference number: A72000; MUID:99206606; PMID:10192388
 A:Accession: G72077
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-316 <ARN>
 A:Cross-references: GB:AE001627; GB:AE001363; NID:G4376721; PIDN:AA018585.1; PID:G437
 A:Experimental source: strain CML029
 C:Genetics:

Query Match 53.9%; Score 41; DB 2; Length 316;
 Best Local Similarity 61.5%; Pred. No. 12;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 AEVSYVHNGAKF 15
 : : : : :
 DB 44 AQVQYLIKVDNAKF 56

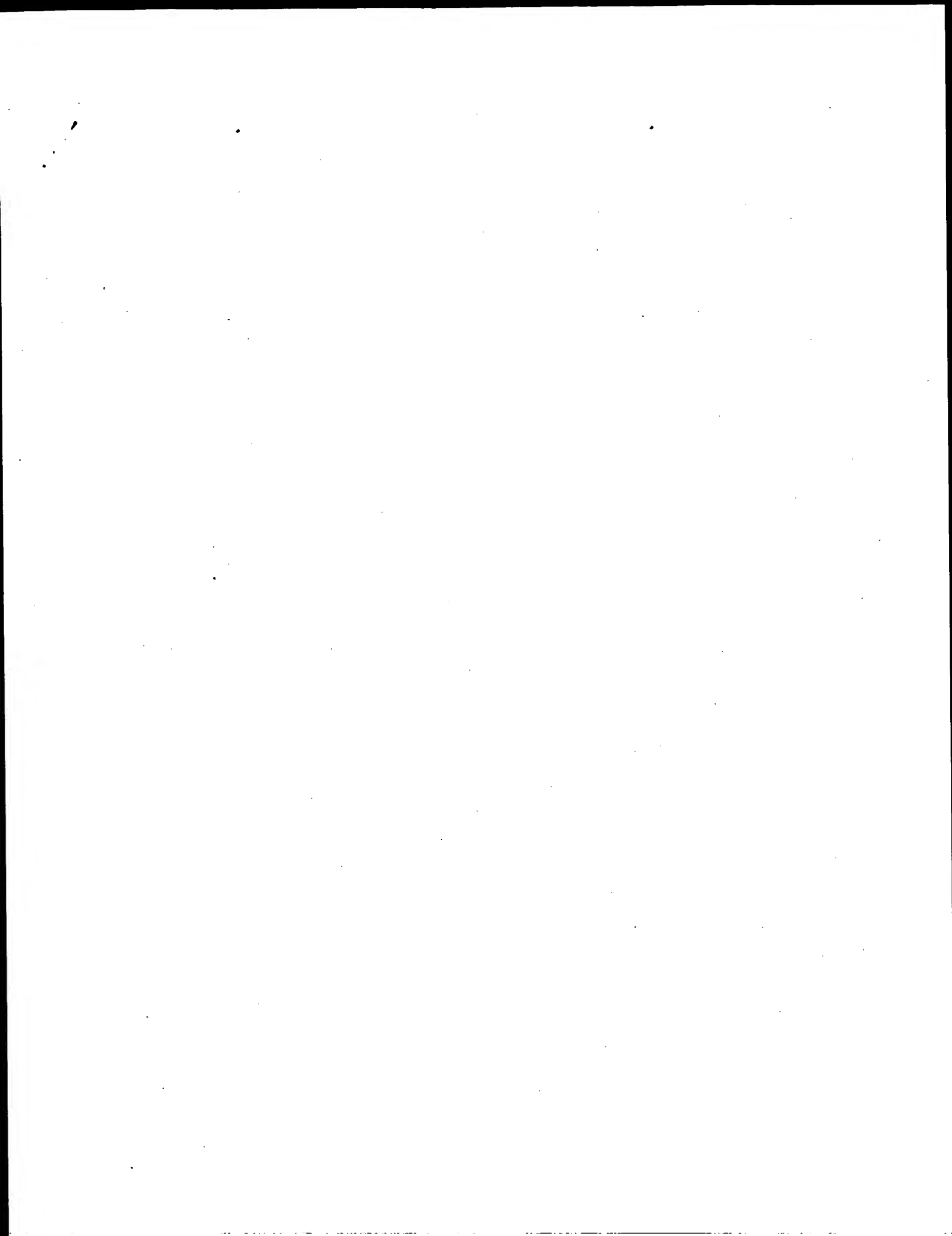
RESULT 15
 G86545
 CT007 hypothetical protein [imported] - *Chlamydomonas reinhardtii* (strain J138)
 C:Species: *Chlamydomonas reinhardtii*
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: G86545
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
 Nucleic Acids Res. 28, 2311-2314, 2000
 A>Title: Comparison of whole genome sequences of *Chlamydia pneumoniae* J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: G86545
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-316 <STO>
 A:Cross-references: GB:BA000008; NID:98978813; PIDN:BA98649.1; GSPDB:GN00142

A:Experimental source: strain J138
C:Genetics:
A:Gene: CPJ0441

Query Match 53.9%; Score 41; DB 2; Length 316;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 AEVSYYHNGAKF 15
|:|:|:|:|:|
Db 44 AQOYLYKVNDAKF 56

Search completed: April 20, 2003, 13:15:57
Job time : 8.07895 secs



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524d-131

Sequence: 1 SRAEVSYHVNAGAF 15

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	514	1	MPA2_CRYJA
2	43	56.6	419	1	CYB_RHOVI
3	42	55.3	385	1	CYB_ACACA
4	42	55.3	386	1	CYB_PICPT
5	41	53.9	316	1	V441_CERPU
6	41	53.9	1307	1	PHYI_CERPU
7	40	52.6	379	1	CYB_IOLBL
8	40	52.6	389	1	CYB_DICDI
9	40	52.6	578	1	YTEM_HAEIN
10	39	51.3	467	1	PGLT_ACTCH
11	38	50.0	134	1	CYB_EORDA
12	38	50.0	348	1	V113_AQUAE
13	38	50.0	385	1	CYB_SACDO
14	38	50.0	385	1	CYB_YEAST
15	38	50.0	386	1	CYB_HANWI
16	38	50.0	423	1	MBI2_YEAST
17	38	50.0	517	1	MBI3_YEAST
18	38	50.0	589	1	VAAL_TREPA
19	38	50.0	589	1	VAAL_TREPA
20	37	48.7	341	1	MURB_PASMU
21	37	48.7	357	1	TRMA_CAMJE
22	37	48.7	379	1	CYB_URSAR
23	37	48.7	379	1	CYB_URSMA
24	37	48.7	380	1	CYB_ONCMW
25	37	48.7	380	1	CYB_SALSA
26	37	48.7	380	1	CYB_SALSA
27	37	48.7	383	1	CYB_SALTR
28	37	48.7	399	1	CYB_PHYME
29	37	48.7	607	1	MODM_RHIVV
30	37	48.7	687	1	KXDV_BACSU
31	37	48.7	733	1	AGAI_PEDPE
32	37	48.7	1385	1	NAC2_CHURE
33	37	48.7	2111	1	YPB4_CABEL

34	36.5	48.0	289	1	YHDF_BACSU
35	36	47.4	79	1	CYB_DIPHR
36	36	47.4	79	1	CYB_DIPHR
37	36	47.4	156	1	GLMS_SPHYA
38	36	47.4	379	1	CYB_NYCCO
39	36	47.4	379	1	CYB_PSENG
40	36	47.4	379	1	CYB_ROMDI
41	36	47.4	392	1	PGK_VIRBCH
42	36	47.4	591	1	GLMS_AQUAE
43	36	47.4	605	1	GLMS_CAUCR
44	36	47.4	607	1	GLMS_AGRTS
45	36	47.4	607	1	NOMI_RHYME

ALIGNMENTS

RESULT 1
MPA2_CRYJA STANDARD; PRT; 514 AA.
AC P43212;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)
DE (Major pollen allergen Cry j 2) (Cry j II).
OS Cyrtomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OX NCBI_TaxId=3369;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=9501077; PubMed=7926035;
RA Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,
RT Usui M., Kurimoto M.,
RT "Molecular cloning of the second major allergen, Cry j II, from
RT Japanese cedar pollen.";
RT FEBS Lett. 353:124-128 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=94271186; PubMed=8002972;
RA Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.,
RT "CDNA cloning and expression of Cry j II the second major allergen of
RT Japanese cedar pollen.";
RL Biochem. Biophys. Res. Commun. 201:1021-1028(1994).
RN [3]
RP SEQUENCE OF 55-64.
RX MEDLINE=90342988; PubMed=2382797;
RA Sakaguchi M., Inouye S., Tanita M., Ando S., Usui M., Matubasi T.,
RT "Identification of the second major allergen of Japanese cedar
RT pollen.";
RL Allergy 45:309-312(1990).
RN [4]
RP CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
RN galactosiduronic linkages in pectate and other galacturonans.
RN - SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).
RN - SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
(POLYGALACTURONASES).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL: D37765; BAA07021.1; -;
DR EMBL: D29772; BAA06172.1; -;
DR HSSP: P26509; 1BHE.
DR InterPro: IPR000743; GH28.
DR Pfam: PF00295; Glyco_hydro_28; 1.

DR PROSITE, PS00502; POLYGALACTURONASE, 1.
 KM Hydrolyase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
 KM Amyloplast; Glycoprotein; Allergen.
 FT SIGNAL 1 22
 FT PROPEP 23 45
 FT CHAIN 46 433
 FT PROPEP 434 514
 FT ACT SITE 278 278
 FT CARBOHYD 460 460
 FT CARBOHYD 472 472
 FT CARBOHYD 5 5
 FT CONFLICT 12 12
 FT CONFLICT 34 35
 FT CONFLICT 37 37
 FT CONFLICT 88 88
 FT CONFLICT 98 98
 FT CONFLICT 451 451
 FT CONFLICT 454 454
 FT CONFLICT 504 504
 FT CONFLICT 507 507
 SQ SEQUENCE 514 AA; 56645 MW; 624611C3FA8D6302 CRC64;

Query Match 100.0%; Score 76; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 2.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAEVSYHVNGAKF 15
 DB 290 SRAEVSYHVNGAKF 304

RESULT 2
 CYB_RHOVI STANDARD; PRT; 419 AA.
 AC P81378;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome B.
 GN PEBB OR FBCH.
 OS Rhodospseudomonas viridis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Hyphomicrobium group; Blastochloris.
 OK NCBI_Taxid=1079;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 133;
 RA MEDLINE=90158506; PubMed=2560136;
 RA Verbit J., Lang F., Gabelini N., Oesterhelt D.;
 RT "Cloning and sequencing of the fbcb, B and C genes encoding the
 cytochrome b/c1 complex from Rhodospseudomonas viridis";
 RL Mol. Gen. Genet. 219:445-452(1989).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 COUPLED TO ATP SYNTHESIS.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CYTOCHROME C1 AND THE RISKE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR InterPro; IPR000179; Cyt b b6.
 DR Pfam; PF000032; cytochrome_b_c1.
 DR Pfam; PF000033; cytochrome_b_N.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR PROSITE; PS00193; CYTOCHROME_B_OO; 1.
 KM Electron transport; Respiratory chain; Heme; Transmembrane.
 FT METAL 96 96
 FT METAL 110 110
 FT METAL 197 197
 FT METAL 211 211
 SQ SEQUENCE 419 AA; 47237 MW; 202C0C3D25B2A6D6 CRC64;

Query Match 56.6%; Score 43; DB 1; Length 419;
 Best Local Similarity 54.5%; Pred. No. 6.7;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 5 VSYVHNGAKF 15
 DB 92 IRYIHNGASF 102

RESULT 3
 CYB_ACACA STANDARD; PRT; 385 AA.
 AC Q37378;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome B.
 GN COB.
 OS Acanthamoeba castellanii (Amoeba).
 OC Mitochondrion.
 OC Eukaryota; Acanthamoebidae; Acanthamoeba.
 OK NCBI_Taxid=5755;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 30010 / NEFF;
 RC MEDLINE=95147275; PubMed=7844823;
 RA Burger G., Plante I., Loneragan K.M., Gray M.W.;
 RT "The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba
 castellanii: complete sequence, gene content and genome
 organization";
 RL J. Mol. Biol. 245:522-537(1995).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 COUPLED TO ATP SYNTHESIS.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CYTOCHROME C1 AND THE RISKE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

CC -----
 DR EMBL; U12386; AAD11830.1; -
 DR InterPro; IPR000179; Cyt b b6.
 DR Pfam; PF000032; cytochrome_b_c1.
 DR Pfam; PF000033; cytochrome_b_N.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR PROSITE; PS00193; CYTOCHROME_B_OO; 1.
 KM Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
 Heme.
 FT METAL 82 82
 FT METAL 96 96
 FT METAL 183 183
 FT METAL 197 197
 SQ SEQUENCE 385 AA; 43934 MW; 5F24F72C431359D5 CRC64;

Query Match 55.3%; Score 42; DB 1; Length 385;
 Best Local Similarity 54.5%; Pred. No. 9.1;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 VSYVHNGAKF 15
 DB 78 IRYIHNGASF 88

RESULT 4

CYB_PICPJ STANDARD; PRT; 386 AA.
 ID CYB_PICPJ Q36507;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome B.
 GN COB OR CYTB.
 OS Pichia pilipeti (Yeast).
 OG Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4928;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN=CBS 2887;
 RC MEDLINE=93204976; PubMed=8455612;
 RA Fukushima H., Sor F., Drissi R., Dinouel N., Miyakawa I., Rousset S.,
 RA Viola A.M.;
 RT "Linear mitochondrial DNAs of yeasts: frequency of occurrence and
 RT general features."
 RL Mol. Cell. Biol. 13:2309-2314(1993).
 CC -1- FUNCTION: COMPONENT OF THE UBIOQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X66593; CAA47155.1; -;
 DR InterPro; IPR00179; Cyt b b6.
 DR Pfam; PF00032; cytochrome_b_c1.
 DR Pfam; PF00033; cytochrome_b_n; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR PROSITE; PS00193; CYTOCHROME_B_OO; FALSE NEG.
 DR Mitochondrion; Electron transport; Respiratory chain; Transmembrane;
 KW Heme.
 FT METAL 82 82 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 96 96 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 183 183 IRON 2 (HEME B562 AXIAL LIGAND).
 FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
 SQ SEQUENCE 386 AA; 43345 MW; CS3A87644963E5B CRC64;
 Query Match 55.3%; Score 42; DB 1; Length 386;
 Best Local Similarity 54.5%; Pred. No. 9.1;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 5 VSYVHVNGAKF 15
 DB 78 IRYIHANGASF 88
 RESULT 5
 ID Y441_CHLPN STANDARD; PRT; 316 AA.
 AC Q928A2; Q9JRW4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein CPN0441/CPJ0441.
 GN CPN0441 OR CPJ0441.
 GN Chlamydia pneumoniae (Chlamydia pneumoniae).

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Klinger L., Greenwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uitterlind T., Berry K., Baes S.,
 RA Linher K., Weidman J., Kfour H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Morn and Chlamydia
 RT pneumoniae AR39."
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kunara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWD029 from USA."
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0441/CT007/TC0275
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001627; AAD18585.1; -;
 DR EMBL; AE002194; AAF73653.1; -;
 DR EMBL; AP002546; BAA98649.1; -;
 DR TIGR; CP0312; -;
 KW Hypothetical protein; Complete proteome.
 FT CONFLICT 243 243 Y -> H (IN REF. 1).
 SQ SEQUENCE 316 AA; 35415 MW; C821CFB53B014B41 CRC64;
 Query Match 53.9%; Score 41; DB 1; Length 316;
 Best Local Similarity 61.5%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 3 AEVSYVHVNGAKF 15
 DB 44 AOVQYLVKVDKXF 56
 RESULT 6
 ID PHYL_CERPU STANDARD; PRT; 1307 AA.
 AC P25848; P93100;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Light-sensor Protein Kinase [includes: Phytochrome; Protein kinase
 DE (EC 2.7.1.-)].
 GN PHYL OR PHY.
 OS Ceratodon purpureus (Moss).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 OC Bryopsida; Dicranidae; Dicranales; Dittichaceae; Ceratodon.
 OX NCBI_TaxID=3225;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9309252; PubMed=1463836;
 RA Thummler F., Dufner M., Kreisel P., Dittich P.;
 RT "Molecular cloning of a novel phytochrome gene of the moss *Ceratodon purpureus* which encodes a putative light-regulated protein kinase.";
 RL Plant Mol. Biol. 20:1003-1017(1992).
 RN [2]
 RP SEQUENCE OF 49-538 FROM N.A.
 RX MEDLINE=91085543; PubMed=2261981;
 RA Thummler F., Beetz A., Ruediger W.;
 RT "Phytochrome in lower plants. Detection and partial sequence of a phytochrome gene in the moss *Ceratodon purpureus* using the polymerase chain reaction.";
 RL FEBS Lett. 275:125-129(1990).
 RN [3]
 RP REVISIONS TO C-TERMINUS.
 RC STRAIN-WT3;
 RA Pasenteis K., Paulo N., Dittich P., Algarra P., Thummler F.;
 RL Dufner M., Kreisel P.;
 CC Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCOVERSION OF PR IN PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-BIPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN, PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: LOCATED IN A FIXED POSITION CLOSE TO THE PLASMA MEMBRANE.
 CC -1- PTM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE PHYTOCHROME FAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U87632; AAB47762.1; -;
 DR EMBL; X17084; CAA34936.1; ALT_SEQ.
 DR PIR; S12966; S12966.
 DR HSR; S27396; S27396.
 DR HSSP; P08631; IAD5.
 DR InterPro; IPR000719; Euk_Pkinase.
 DR InterPro; IPR003018; GAF.
 DR InterPro; IPR000700; PAS-associ_C.
 DR InterPro; IPR000014; PAS domain.
 DR InterPro; IPR001294; Phytochrome.
 DR InterPro; IPR004040; STY_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; Pkinase_1.
 DR Pfam; PF00360; Phytochrome; 1.
 DR Pfam; PF00989; PAS; 1.
 DR Pfam; PF01590; GAF; 1.
 DR PRINTS; PR01033; PHYTOCHROME.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00065; GAF; 1.
 DR SMART; SM00091; PAS; 1.
 DR SMART; SM00221; STYK; 1.
 DR TIGRfam; TIGR00229; sensory_box; 1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00245; PHYTOCHROME_1; 1.
 DR PROSITE; PS00046; PHYTOCHROME_2; 1.
 DR PROSITE; PS00112; PAS; 1.
 DR PROSITE; PS00113; PAC; 1.
 DR Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 609 680
 FT DOMAIN 683 739
 FT DOMAIN 779 1003
 FT DOMAIN 1004 1307
 FT BINDING 320 320
 FT BINDING 1010 1018
 FT NP_BIND 1031 1031
 FT BINDING 1127 1127
 FT ACT_SITE 1127 1127
 SQ SEQUENCE 1307 AA; 145843 MW; ESE77A9FEF301A5C CRC64;
 Query Match 53.9%; Score 41; DB 1; Length 1307;
 Best Local Similarity 46.7%; Pred. No. 41;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Db 1141 TRSRIQVHVYKADF 1155
 QY 1 SRAEVSIVYVNGAKE 15
 DB ::::|::|::|
 ID CYB_L0LBL STANDARD; PRT; 379 AA.
 CYB_L0LBL
 AC 047477;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome B.
 GN COB OR CYTB.
 OS Loligo bleekeri (Bleeker's squid).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Teuthoidea;
 OC Myxopsida; Loliginidae; Loligo.
 OX NCBI_Taxid=6617;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tomita K., Ueda T., Watanabe K.;
 RT "Completing of squid (*Loligo bleekeri*) mitochondrial genome sequencing.";
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL COUPLED TO ATP SYNTHESIS.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B, CYTOCHROME C1 AND THE RIESKE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB009838; BAA24060.1; -;
 DR EMBL; AB029616; BAB03647.1; -;
 DR InterPro; IPR000179; Cyt_b_b6.
 DR Pfam; PF00032; cytochrome_b_c1; 1.
 DR Pfam; PF00033; cytochrome_b_n; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR PROSITE; PS00193; CYTOCHROME_B_CO; 1.

KM Mitochondrion; Electron transport; Respiratory chain; Transmembrane;
Heme.
FT METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 183 183 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 379 AA; 43399 MW; DFFFLA057B33CA CRC64;

Query Match 52.6%; Score 40; DB 1; Length 379;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 YHVNGAKF 15
|:|:|:|:|
DB 82 YIHANGASF 90

RESULT 8
CYB_DICDI STANDARD; PRT; 389 AA.
AC Q37311;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Dictyostelium discoideum (slime mold).
OG Mitochondrion.
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AX3;
RX MEDLINE=95254668; PubMed=7736610;
RA Angata K., Kuroe K., Yanagisawa K., Tanaka Y.;
RT "Codon usage, genetic code and phylogeny of Dictyostelium discoideum
mitochondrial DNA as deduced from a 7.3-kb region.";
RL Curr. Genet. 27:249-256(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=AX3;
RX MEDLINE=20279206; PubMed=10821186;
RA Ogawa S., Yoshino R., Angata K., Iwamoto M., Pi M., Kuroe K.,
RA Matsuo K., Morio T., Urushihara H., Yanagisawa K., Tanaka Y.;
RT "The mitochondrial DNA of Dictyostelium discoideum: complete sequence,
gene content and genome organization.";
RL Mol. Genet. 263:514-519(2000).
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
COUPLED TO ATP SYNTHESIS.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; D1466; BAA03933.1; -
DR EMBL; AB000109; BAA78061.1; -
DR DickeyDB: DD07272; cob.
DR InterPro: IPR000179; Cyt_b_b6.
DR Pfam: PF00032; cytochrome_b_c1.
DR Pfam: PF00033; cytochrome_b_n; 1.
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE: PS00193; CYTOCHROME_B_OO; 1.

KM Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
Heme.
FT METAL 82 82 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 96 96 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 183 183 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 389 AA; 44483 MW; PFF9E7092AAEBB1 CRC64;

Query Match 52.6%; Score 40; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 YHVNGAKF 15
|:|:|:|:|
DB 80 YIHANGASF 88

RESULT 9
YTFM_HAEIN STANDARD; PRT; 578 AA.
AC P44038;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein H10698 precursor.
GN H10698.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=9530630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs C.A., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
RN [2]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae";
RL Electrophoresis 21:411-429(2000).
CC -1- SIMILARITY: STRONG, TO E. COLI YTFM.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U32752; AAC22357.1; -
DR TIGR: H10698;
DR InterPro: IPR000184; Bac_surfing_D15.
DR Pfam: PF01103; Bac_surface_Ag; 1.
DR Signal: Complete proteome.
DR SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 578 PROTEIN H10698.
SQ SEQUENCE 578 AA; 65012 MW; 34F9AC189C505876 CRC64;

Query Match 52.6%; Score 40; DB 1; Length 578;

Best Local Similarity 50.0%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 1 SRAEVSYYHNGAK 12
DB 447 ABAEIGYHRTKG 458

RESULT 10

PGLR ACTCH STANDARD; PRT; 467 AA.
AC P3536;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Polylacturonase precursor (EC 3.2.1.15) (PG) (Pectinase).
OS Actinidia chinensis (Kiwi) (Yangtze).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Ericales; Actinidiaceae; Actinidia.
OX NCBI_TaxID=3625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Deliciosa;
RX MEDLINE=94302157; PubMed=8029342;
RA Atkinson R.G., Gardner R.C.;
RT "A polylacturonase gene from kiwifruit (Actinidia deliciosa).";
RL Plant Physiol. 103:669-670 (1993).
CC -1- FUNCTION: ACTS IN CONCERT WITH THE PECTINESTERASE, IN THE RIPENING
PROCESS. IS INVOLVED IN CELL WALL METABOLISM, SPECIFICALLY IN
POLYURONIDE DEGRADATION.
CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
galactosiduronic linkages in pectate and other galacturonans.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DEVELOPMENTAL STAGE: IN RIPENING FRUIT.
CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
(POLYGLACTURONASES).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; L12019; AAC14453.1; -
DR InterPro: IPR000743; GH28.
DR Pfam: PF00295; Glyco_hydro_28; 1.
DR PROSITE: PS00502; POLYGLACTURONASE; 1.
KW Hydroxylase; Glycosidase; Cell wall; Signal; Fruit ripening;
KW Glycoprotein.
FT SIGNAL 1 27
FT CHAIN 28 467 POTENTIAL.
FT ACT SITE 306 306 POLYGLACTURONASE.
FT CARBOHYD 290 290 PROBABLE.
SQ SEQUENCE 467 AA; 50776 MW; 5A9A61483C028B7A CRC64;

Query Match 51.3%; Score 39; DB 1; Length 467;
Best Local Similarity 71.4%; Pred. No. 34;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 SRAEVSYYHNGAK 14
DB 318 SEAHVSQVYNGAK 331

RESULT 11

CYB_SORDA STANDARD; PRT; 134 AA.
AC O21411; O21410;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome B (Fragment).
GN MTCYB OR COB OR CYTB.
OS Sorex daphaenodon (Large-toothed shrew).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Soricidae; Soricinae; Sorex.
OX NCBI_TaxID=62272;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=#696, and 95/8/19-1; TISSUE=Hindfoot;
RA Ondaachi S., Masuda R., Abe H., Adachi J., Dokuchaev N.E.,
RA Hanksalini V., Yoshida M.C.;
RT "Molecular phylogeny from nucleotide sequences of the mitochondrial
cytochrome b gene and evolutionary history of Eurasian soricine shrews
(Mammalia, Insectivora).";
RL Zool. Sci. 14:527-532 (1997).
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
COUPLED TO ATP SYNTHESIS.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CYTOCHROME C1 AND THE RISKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; D85356; BAA21349.1; -
DR EMBL; D85342; BAA21335.1; -
DR InterPro: IPR001079; Cyt b b6.
DR Pfam: PF00033; cytochrome b N; 1.
DR PROSITE: PS00193; CYTOCHROME_B_HEME; 1.
DR PROSITE: PS00193; CYTOCHROME_B_OO; PARTIAL.
KW Mitochondrion; Electron transport; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 83 83 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 97 97 IRON 2 (HEME B566 AXIAL LIGAND).
FT VARIANT 82 84 LHV -> FHA (IN 95/8/19-1).
FT NON TER 134 134
SQ SEQUENCE 134 AA; 15218 MW; 3D53D624E81B7F7C CRC64;

Query Match 50.0%; Score 38; DB 1; Length 134;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 5 VSYVHNGA 13
DB 79 IRYLHVNGA 87

RESULT 12

Y113_AQUAE STANDARD; PRT; 348 AA.
AC O66516;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AQ_113 precursor.
GN AQ_113.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-VF5;
RA DECKERT G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Gidman D.E., Overbeek R., Sneed M.A., Keller M., Anujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aerolicus".
RL Nature 392:353-358 (1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A000674; AAC06480.1; -
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 348 HYPOTHETICAL PROTEIN A0.113.
SQ SEQUENCE 348 AA; 38760 MW; BDDC09E0013AF152 CRC64;
Query Match 50.0%; Score 38; DB 1; Length 348;
Best Local Similarity 57.1%; Pred. No. 38;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 RAEVYVHNGAKF 15
DB 316 RAEVYVNTDNKVF 329
RESULT 13
CYB_SACDO STANDARD; PRT; 385 AA.
ID CYB_SACDO 035819;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Saccharomyces douglasii (Yeast).
OG Mitochondrion.
OC Baktaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxId=46617;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91218158; PubMed=1708831;
RA Tian G.L., Michel F., Macadre C., Slonimski P.P., Lazowska J.;
RT "Incipient mitochondrial evolution in yeasts. II. The complete
RT sequence of the gene coding for cytochrome b in Saccharomycetes
RT douglasii reveals the presence of both new and conserved introns and
RT discloses major differences in the fixation of mutations in
RT evolution."
RL J. Mol. Biol. 218:747-760 (1991).
CC -----
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RISKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC -----
DR EMBL; X59280; CAA41971.1; -
DR InterPro: IPR000179; Cyt_b_b6.
DR Pfam: PF00032; cytochrome_b_c1.
DR Pfam: PF00033; cytochrome_b_n; 1.
DR PROSITE: PS00192; CYTOCHROME B HEME; 1.
DR PROSITE: PS00193; CYTOCHROME B OO; 1.
KW Mitochondrion; Electron transport; Respiratory chain; Transmembrane;
Heme.
FT METAL 82 82 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 96 96 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 183 183 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 385 AA; 43631 MW; 3541B2C7D77B2DD2 CRC64;
Query Match 50.0%; Score 38; DB 1; Length 385;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 YVHNGAKF 15
DB 80 YVHNGASF 88
RESULT 14
CYB_YEAST STANDARD; PRT; 385 AA.
ID CYB_YEAST P00163; Q35807; Q36301; Q35802;
AC P00163; Q35807; Q36301; Q35802;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Saccharomyces cerevisiae (Baker's yeast).
OG Mitochondrion.
OC Baktaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=D273-10B/A21;
RA MEDLINE=81046788; PubMed=6253454;
RA Nobrega F.G., Tzagoloff A.;
RT "Assembly of the mitochondrial membrane system. DNA sequence and
RT organization of the cytochrome b gene in Saccharomyces cerevisiae
RT D273-10B".
RL J. Biol. Chem. 255:9828-9837 (1980).
CC -----
CC [2]
RP SEQUENCE FROM N.A.
RA STRAIN=D273-10B/A21;
RA MEDLINE=85000973; PubMed=6383504;
RA Bonnard C.A., Nobrega F.G.;
RT "Revision of the nucleotide sequence at the last intron of the
RT mitochondrial apocytochrome b gene in Saccharomyces cerevisiae".
RL Braz. J. Med. Biol. Res. 17:17-20 (1984).
CC -----
CC [3]
RP SEQUENCE FROM N.A.
RA STRAIN=MR200;
RA MEDLINE=95255283; PubMed=7737175;
RA Claes M.G., Pera J., Shu Y., Samatey F.A., Popot J.L., Jacq C.;
RT "Mutations in vivo import of hydrophobic proteins into yeast
RT mitochondria. The case of a cytoplasmically synthesized apocytochrome
RT b".
RL Eur. J. Biochem. 228:762-771 (1995).
CC -----
CC [4]
RP SEQUENCE OF 20-143 FROM N.A.
RA STRAIN=777-3A;
RA MEDLINE=81088336; PubMed=7004642;
RA Lazowska J., Jacq C., Slonimski P.P.;
RT "Sequence of introns and flanking exons in wild-type and box3 mutants
RT of cytochrome b reveals an interlaced splicing protein coded by an
RT intron".
RL Cell 22:353-348 (1980).
CC -----


```

RN [5]
RP SEQUENCE OF 144-169 FROM N.A.
RC STRAIN=777-3A;
RX MEDLINE=82115326; PubMed=7034963;
RA Lazowska J., Jacq C., Slonimski P.P.;
RT "Splice point of the third intron in the yeast mitochondrial
   cytochrome b gene.";
RL Cell 27:12-14 (1981).
RN [6]
RP MUTANT W7.
RX MEDLINE=90005972; PubMed=2551711;
RA Brivet-Chevillotte P., di Rago J.-P.;
RT "Electron-transfer restoration by vitamin K3 in a complex
   III-deficient mutant of S. cerevisiae and sequence of the
   corresponding cytochrome b mutation.";
RL FEBS Lett. 255:5-9 (1989).
CC -1- FUNCTION: COMPONENT OF THE UBIOQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01476; AAA99924.1; ALT_SEQ.
DR EMBL; X84042; CAA58861.1; -.
DR EMBL; V00686; -; NOT ANNOTATED_CDS.
DR EMBL; J01473; AAA32151.2; -.
DR EMBL; J01472; AAA32151.2; JOINED.
DR EMBL; J01475; AAA32152.2; -.
DR EMBL; J01474; AAA32152.2; JOINED.
DR PIR; A00159; CBBY.
DR SGD; S0007270; COB.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_OO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 82 82 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 96 96 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 183 183 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
FT VARIANT 131 131 G -> S (IN MUTANT W7 WHICH IS RESPIRATORY
FT DEFICIENT).
FT CONFLICT 122 122 T -> I (IN REF. 1 AND 2).
FT CONFLICT 270 270 D -> V (IN REF. 3).
SQ SEQUENCE 385 AA; 43659 MW; 23C41614B712A79C CRC64;

Query Match 50.0%; Score 38; DB 1; Length 385;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 YVHVGAKF 15
DB 80 YLHANGASF 88

```

```

DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Hansenula wingei (Yeast).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4907;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=21;
RA Sekito T., Okamoto K., Kitano H., Yoshida K.;
RT "Yeast Hansenula wingei mitochondria genome's complete DNA sequence
   demonstrated unique characteristics.";
RL Nucleic Acids Symp. Ser. 31:233-234 (1994).
CC -1- FUNCTION: COMPONENT OF THE UBIOQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D31785; BAA06572.1; -.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_OO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 82 82 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 96 96 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 183 183 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 386 AA; 44083 MW; 8B46CD4553C643B8 CRC64;

Query Match 50.0%; Score 38; DB 1; Length 386;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 YVHVGAKF 15
DB 80 YLHANGASF 88

```

Search completed: April 20, 2003, 13:07:45
 Job time : 3.92105 secs

RESULT 15
 CYB_HANNI
 ID CYB_HANNI
 AC P48877; STANDARD; PRT; 386 AA.

Query Match	80.3%;	Score 61;	DB 10;	Length 507;
Best Local Similarity	73.3%;	Pred. No. 0.0051;		
Matches 11; Conservative	4;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 SRAEVSYVHNGAKF 15

Db 291 SRSEVSVHLDGAKF 305

RESULT 2

Q8UG30 PRELIMINARY; PRT; 155 AA.
AC Q8UG30;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Hypothetical protein Atu1212.
GN Atu1212 OR AGC_2237.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970);
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Bisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland B., Palmeri A., Gordon D.,
RA Raymond C., Rouse G., Saenphitachak C., Wu Z., Romero P., Krespan W., Perry M.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeser B.W.,
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58.";
RT Science 294:2317-2323 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194; Miller N., Blanchard M.,
RA Goodner B., Hinkle G., Gattung S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Quicilo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Hountel K., Gordon J., Vaudin M., Tarchnouk O., Bpp A., Liu F.,
RA Houliam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58.";
RT Science 294:2323-2328 (2001).
DR EMBL; AF009084; AAL42223.1; -;
DR EMBL; AF009085; AAK87014.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 155 AA; 16849 MW; C04F0F4540F0E67 CRC64;

Query Match 59.2%; Score 45; DB 16; Length 155;
Best Local Similarity 61.5%; Pred. No. 1.5;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 RAESVYVHNGAK 14
Db 132 RADVAVHVSAR 144

RESULT 3

O23147 PRELIMINARY; PRT; 431 AA.
AC O23147;
DT 01-JUN-1998 (TREMBlrel. 05, Created)
DT 01-JUN-1998 (TREMBlrel. 05, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Endo-polygalacturonase.
GN PGAL OR ADPGI OR T8H10.110.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RA Sander L., Child R., Ulyskov P., Albrechtsen M., Joergensen B.,
RA Borkhardt B.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERRECTA;
RA Jenkins E.S., Roberts J.A.;
RT "deniscence-related expression of an Arabidopsis thaliana gene
encoding a polygalacturonase in transgenic plants of Brassica napus.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.
RA Benes V., Reclmann S., Borkova D., Anseorge W., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Queller F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
(POLYGALACTURONASES).
DR EMBL; AJ002532; CAA05525.1; -;
DR EMBL; AF037367; AAC98923.1; -;
DR EMBL; AL133248; CAB6108.1; -;
DR InterPro; IPR000743; GH28
DR InterPro; IPR004048; Reg_chir_condens.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
DR PROSITE; PS00626; RCCL 2; UNKNOWN 1.
KW Cell wall; Glycosidase; Hydrolase; Signal.
SQ SEQUENCE 431 AA; 46572 MW; 0F0F30DF45804FE7 CRC64;

Query Match 56.6%; Score 43; DB 10; Length 431;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAVSVHNGAKF 15
Db 299 SKAVSVHNGAKF 313

RESULT 4

O29131 PRELIMINARY; PRT; 456 AA.
AC O29131;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Hypothetical protein AF1134.
GN AF1134.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=96049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson U.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrtides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kitzness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.T., McNeill L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weiman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-

RT reducing archaeon *Archaeoglobus fulgidus*;
 RL Nature 390:364-370(1997).
 DR EMBL: AB001026; MAB90123.1; -
 DR TIGR: AF1134; -
 DR InterPro: IPR000379; Ser esters site.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 456 AA; 51748 MW; 37CB62B58CC9357 CRC64;

Query Match 56.6%; Score 43; DB 17; Length 456;
 Best Local Similarity 72.7%; Pred. No. 12;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 VSVYVNGAKF 15
 |||||
 DB 178 VSVYVNGAKF 188

RESULT 5

ID 036562 PRELIMINARY; PRT; 131 AA.

AC 036562; (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 GN Apocytochrome B (Fragment).
 OS Cytb.
 OS Williopsis mrakii (Yeast) (Hansenula mrakii).
 OC Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Williopsis.
 NCBI_TaxID=4963;
 [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=CBS1707;
 RX MEDLINE=93204976; PubMed=8455612;
 RA Fukuhara H., Sor F., Drissi R., Dinouel N., Miyakawa I., Rousset S.,
 RA Viola A.M.;
 RT "Linear mitochondrial DNAs of yeast. I. Occurrence and general
 RT features";
 RL Mol. Cell. Biol. 13:2309-2314(1993).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=CBS1707;
 RX MEDLINE=94287716; PubMed=8017108;
 RA Drissi R., Sor F., Fukuhara H.;
 RT "Genes of the linear mitochondrial DNA of Williopsis mrakii: Coding
 RT sequences for a maturase-like protein, a ribosomal protein VAR1
 RT homologue, cytochrome oxidase subunit 2 and methyl tRNA";
 RL Yeast 10:391-398(1994).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).

CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RISKE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

DR EMBL: X66594; CA47157.1; -
 DR InterPro: IPR000179; Cyt_b_b6.
 DR Pfam: PF00033; cytochrome_b_N; 1.
 DR PROSITE: PS00192; CYTOCHROME B HEME; 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain;
 KW Transmembrane;
 FT NON_TER
 SQ SEQUENCE 131 AA; 14799 MW; 835A67A6B8B23265 CRC64;

Query Match 55.3%; Score 42; DB 8; Length 131;
 Best Local Similarity 54.5%; Pred. No. 4.5;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 VSVYVNGAKF 15
 :|||

DB 78 IRYIHANGASF 88

RESULT 6

ID 047792 PRELIMINARY; PRT; 380 AA.

AC 047792;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE Cytochrome b (Fragment).
 GN CYTB.
 OS Xenocypis yumanensis.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Xenocypis.
 NCBI_TaxID=70549;
 [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=MUSCLE;
 RX MEDLINE=21097357; PubMed=11161753;
 RA Xiao W., Zhang Y., Liu H.;
 RT "Molecular Systematics of Xenocypinae (Teleostei: Cyprinidae):
 RT Taxonomy, Biogeography, and Coevolution of a Special Group Restricted
 RT in East Asia";
 RL Mol. Phylogenet. Evol. 18:163-173(2001).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).

CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RISKE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

DR EMBL: AF036208; AAC15610.1; -
 DR InterPro: IPR000179; Cyt_b_b6.
 DR Pfam: PF00032; cytochrome_b_C; 1.
 DR Pfam: PF00033; cytochrome_b_N; 1.
 DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
 DR PROSITE: PS00193; CYTOCHROME_B_OO; UNKNOWN 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain;
 KW Transmembrane.
 FT NON_TER
 SQ SEQUENCE 380 AA; 42988 MW; E5CDD7E9E93D42CF CRC64;

Query Match 55.3%; Score 42; DB 8; Length 380;
 Best Local Similarity 54.5%; Pred. No. 15;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 VSVYVNGAKF 15
 :|||

DB 79 IRYIHANGASF 89

RESULT 7

ID 09684 PRELIMINARY; PRT; 497 AA.

AC 09684;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Apocytochrome b (EC 1.10.2.2).
 GN CYTB OR COB.
 OS Naegleria gruberi.
 OC Mitochondrion.

OC Eukaryota; Eukaryota; Schizozymenida; Vahlkampfiidae; Naegleria.
 OC NCBI_TaxID=5762;
 [1]

RP SEQUENCE FROM N.A.
 RA Burger G., Lang B.F., Neraud T.A., Gray M.W.;
 RT "The mitochondrial genome of the supposedly primitive protist,

```

RT Naegleria gruberi." to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY). WHICH ARE NOT COVALENTLY
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF288092; AAG17790.1; -.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c1; 1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_CO; UNKNOWN; 1.
DR Electon transport; Heme; Mitochondrion; Oxidoreductase;
KM Respiratory chain; Transmembrane.
SQ SEQUENCE 497 AA; 58323 MW; 7A77A3FEE8F690E2 CRC64;

Query Match 55.3%; Score 42; DB 8; Length 497;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 VSYVHNGAKF 15
: |||
Db 82 IRYIHNGASF 92

RESULT 8
ID 054852 PRELIMINARY; PRT; 1195 AA.
AC 054852;
DT 01-JUN-1998 (TREMBLrel. 06, Created).
DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Potassium channel.
GN ERG3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98054206; PubMed=9390998;
RA Shi W., Wymore R.S., Wang H.S., Pan Z., Cohen I.S., McKinnon D.,
RA Dixon J.E.;
RT "Identification of two nervous system-specific members of the erg
RT potassium channel gene family."
RL J. Neurosci. 17:9423-9432(1997).
DR EMBL; AF016191; AAB94741.1; -.
DR InterPro; IPR000595; cNMP_binding.
DR InterPro; IPR003967; Erg_channel.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR000636; M+channel_nlg.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00027; cNMP_binding; 1.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF00785; PAC; 1.
DR PRINTS; PRO1470; ERGCHANNEL.
DR SMART; SM00100; cNMP; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 1.
DR PROSITE; PSS0042; cNMP_BINDING_3; 1.
SQ SEQUENCE 1195 AA; 134900 MW; BA24C54BB86C59B7 CRC64;

Query Match 55.3%; Score 42; DB 11; Length 1195;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RAESYVHNGAKF 15
: |||

```

```

Db 93 KVEVYYHNGSTF 106

RESULT 9
ID 09ER47 PRELIMINARY; PRT; 1195 AA.
AC 09ER47;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Erg3 protein.
GN ERG3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Arcangelii A.;
RT "Erg genes expression during development of mouse embryos."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; A0291608; CAC14797.1; -.
DR InterPro; IPR000595; cNMP_binding.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR000636; M+channel_nlg.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00027; cNMP_binding; 1.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF00785; PAC; 1.
DR SMART; SM00100; cNMP; 1.
DR SMART; SM00086; PAC; 1.
DR PROSITE; PSS0042; cNMP_BINDING_3; 1.
SQ SEQUENCE 1195 AA; 135026 MW; 8DCCB9BA0580FFC2 CRC64;

Query Match 55.3%; Score 42; DB 11; Length 1195;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RAESYVHNGAKF 15
: |||
Db 93 KVEVYYHNGSTF 106

RESULT 10
ID 09NS40 PRELIMINARY; PRT; 1196 AA.
AC 09NS40;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Potassium channel subunit.
GN HERG-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Ganetzky B., Titus S.A.;
RT "Polynucleotides encoding herg-3 potassium channel."
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF032897; AAD01946.1; -.
DR InterPro; IPR000595; cNMP_binding.
DR InterPro; IPR003967; Erg_channel.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR000636; M+channel_nlg.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00027; cNMP_binding; 1.
DR Pfam; PF00520; Ion_trans; 1.

```

DR Pfam: PF00785; PAC; 1.
 DR PRINTS: PR01470; ERCHANNEL.
 DR SMART: SM00100; CNMP; 1.
 DR SMART: SM00086; PAC; 1.
 DR SMART: SM00091; PAC; 1.
 DR PROSITE: PS50042; CNMP BINDING 3; 1.
 SQ SEQUENCE 1196 AA; 135012 MW; 7CE10C1A17EDCAF6 CRC64;

Query Match 55.3%; Score 42; DB 4; Length 1196;
 Best Local Similarity 50.0%; Pred. No. 54;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RAESYVHNGAKF 15
 DB 93 KVEVTVYHKGSTF 106

RESULT 11
 Q97VZ6 PRELIMINARY; PRT; 198 AA.

AC Q97VZ6; 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DE Endonuclease V (nfi) (EC 3.1.25.1).
 GN NFI OR SSO2454
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=2132296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aways M.J., Chan-Weiner C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moers A., Erusio G., Fletcher C., Gordon P.M.K.,
 RA Heimke de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL: AE006844; AKA42594.1;
 KW Hydrolyase; Complete proteome.
 SQ SEQUENCE 198 AA; 22051 MW; DE752959A649398B CRC64;

Query Match 53.9%; Score 41; DB 17; Length 198;
 Best Local Similarity 42.9%; Pred. No. 11;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAEVSIVHNGAK 14
 DB 139 NESSEIVYVINGEK 152

RESULT 12
 Q9HUZ3 PRELIMINARY; PRT; 563 AA.

AC Q9HUZ3; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE Hypothetical protein Ta0818.
 GN TA0818.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;

RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Fishman D., Stocker S., Lups A.N., Baumeister W.;
 RT "The genome sequence of the thermophilic scavenger Thermoplasma
 RT acidophilum.";
 RL Nature 407:508-513 (2000).
 DR EMBL: AL445065; CAC11947.1;
 DR InterPro: IPR004165; CoA_trans.
 DR Pfam: PF01144; CoA_trans; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 563 AA; 62138 MW; A144326E0D7129F5 CRC64;

Query Match 53.9%; Score 41; DB 17; Length 563;
 Best Local Similarity 50.0%; Pred. No. 36;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 RAESYVHNGAKF 15
 DB 165 RANVSLVHIDEEY 178

RESULT 13
 Q8ZSM4 PRELIMINARY; PRT; 608 AA.

AC Q8ZSM4; 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DE Protease IV, conjectural.
 GN PAB3679.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX PubMed=11792869;
 RA Filt-Gibson S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 RT aerophilum.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL: AB009945; AAL65089.1;
 DR InterPro: IPR002142; Peptidase_U7.
 DR Pfam: PF01343; Peptidase_U7; 1.
 DR ProDom: PD002897; Peptidase_U7; 1.
 KW Complete proteome.
 SQ SEQUENCE 608 AA; 66946 MW; F7357A6501ADCE61 CRC64;

Query Match 53.9%; Score 41; DB 17; Length 608;
 Best Local Similarity 61.5%; Pred. No. 39;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 AEVSIVHNGAKF 15
 DB 492 ANMOYVHNGSLF 504

RESULT 14
 P93098 PRELIMINARY; PRT; 1299 AA.

AC P93098; 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DE Ceratodon purpureus (Moss).
 GN Ceratodon purpureus (Moss).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 OC Bryopsida; Dicranidae; Dicranales; Dicranaceae; Ceratodon.
 OX NCBI_TaxID=3225;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WT3;

RA Pasentsis K., Paulo N., Dittich P., Algarra P., Thuenmler F.;
 RT "Characterization and expression of the phytochrome gene family in the
 RT moss *Ceratodon purpureus* (Hedw.) Brid.";
 RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: BELONGS TO THE GNR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.

DR EMBL; Y10901; CAA71838.1; -.
 DR HSSP; P08631; IAD5.
 DR InterPro; IPR002106; ALCRNA_ligase1.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003018; GAF.
 DR InterPro; IPR000524; HTH_Gntr.
 DR InterPro; IPR000700; PAS-asso_C.
 DR InterPro; IPR000014; PAS_domain.
 DR InterPro; IPR001294; Phytochrome.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR Pfam; PF01590; GAF; 1.
 DR Pfam; PF00989; PAS; 1.
 DR Pfam; PF00360; phytochrome; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00035; HTGNT.
 DR PRINTS; PR01033; PHYTOCHROME.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00065; GAF; 1.
 DR SMART; SM00091; PAS; 1.
 DR SMART; SM00221; STYK; 1.
 DR TIGRFAMs; TIGR00229; sensory_box; 1.
 DR PROSITE; PS00339; AA_TRNA_LIGASE_11_2; UNKNOWN_1.
 DR PROSITE; PS00245; PHYTOCHROME_1; 1.
 DR PROSITE; PS50046; PHYTOCHROME_2; 1.
 DR PROSITE; PS500107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW DNA-binding; Receptor; Serine/threonine-protein kinase;
 KW Transcription regulation.
 FT NON_TER 1 1
 SQ SEQUENCE 1299 AA; 144706 MW; 85BDFC2A63212EFA CRC64;

Query Match 53.9%; Score 41; DB 10; Length 1299;
 Best Local Similarity 46.7%; Pred. No. 92;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRAVSVYHVGAKF 15
 DB 1133 TKSEIGYHVXVADF 1147

RESULT 15
 OQ8AB4 PRELIMINARY; PRT; 82 AA.
 AC OQ8AB4;
 DT 01-JUN-2002 (TREMREL. 21, Created)
 DT 01-JUN-2002 (TREMREL. 21, Last sequence update)
 DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
 DE Polyalacturonase (EC 3.2.1.15) (Fragment).
 OS Capsicum frutescens (chili pepper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
 OX NCBI_TaxID=4073;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rao G.U., Paran I.;
 RT "Polygalacturonase: A candidate gene for the soft flesh and deciduous
 RT fruit mutation in Capsicum."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF465938; AAL76254.1; -.
 KW Hydrolyase; Glycosidase.
 FT NON_TER 1 1
 SQ SEQUENCE 82 AA; 8364 MW; F2ECE39B2C4CEB92 CRC64;

Query Match 52.6%; Score 40; DB 10; Length 82;
 Best Local Similarity 60.0%; Pred. No. 6.3; 4; Indels 0; Gaps 0;
 Matches 9; Conservative 2; Mismatches 2;
 OY 1 SRAVSVYHVGAKF 15
 DB 54 SEAHVSDINVYAKF 68

Search completed: April 20, 2003, 13:13:21
 Job time: 15.6711 secs

GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 6.07895 Seconds
(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524D-152

Sequence: 1 LSDISLKLTSKTIAS 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	100.0	514	2 S48730	Cry j II protein -
2	67	100.0	514	2 JC2498	second major aller
3	54	80.6	514	2 JC7100	polygalacturonase
4	51	76.1	507	2 JC7366	Jun a 2 protein -
5	45	67.2	251	2 AH0742	high-affinity zinc
6	44	65.7	242	2 P95256	ABC transporter, A
7	44	65.7	242	2 G98121	hypothetical prote
8	43	64.2	251	2 B64948	probable ABC trans
9	43	64.2	251	2 D85798	hypothetical prote
10	43	64.2	251	2 H90949	probable ABC trans
11	42	62.7	255	1 ORECM3	membrane-bound iro
12	40	59.7	133	2 A12834	conserved hypotet
13	40	59.7	134	2 D97612	hypothetical 14.2K
14	40	59.7	253	2 AD0251	high-affinity zinc
15	40	59.7	253	2 H90033	hypothetical prote
16	40	59.7	317	2 T39736	hypothetical prote
17	39	58.2	93	2 H83968	hypothetical prote
18	39	58.2	292	2 AG0791	hypothetical prote
19	39	58.2	335	2 T04273	probable transcrip
20	39	58.2	839	2 F85334	hypothetical prote
21	39	58.2	1446	2 T04548	myosin heavy chain
22	38	56.7	226	2 AB1662	ABC transporter, A
23	38	56.7	226	2 AD1290	ABC transporter, A
24	38	56.7	230	2 F83697	ABC transporter, A
25	38	56.7	241	2 F82065	ABC transporter, A
26	38	56.7	251	2 G81293	enterochelin uptak
27	38	56.7	273	2 H82286	phosphate ABC trans
28	38	56.7	231	2 AB1025	atac family regula
29	38	56.7	316	2 T33180	hypothetical prote

30	38	56.7	398	2 C84780	hypothetical prote
31	38	56.7	505	2 C90569	hypothetical prote
32	38	56.7	529	2 T42584	legume protein 4
33	38	56.7	530	1 WZBED4	gene 40 protein -
34	38	56.7	548	2 T47548	hypothetical prote
35	37	55.2	142	2 H69745	hypothetical prote
36	37	55.2	165	2 F89835	hypothetical prote
37	37	55.2	231	2 C70000	hypothetical prote
38	37	55.2	232	2 C69822	ABC transporter (A
39	37	55.2	240	2 G71905	glycine betaine/L-
40	37	55.2	240	2 C64609	probable ABC trans
41	37	55.2	253	2 H89845	ABC transporter, A
42	37	55.2	254	2 A95410	hypothetical prote
43	37	55.2	255	2 AC0023	probable ABC trans
44	37	55.2	269	2 D75631	probable taurine t
45	37	55.2	283	2 T44799	iron ABC transport

ALIGNMENTS

RESULT 1

S48730 Cry j II protein - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)

C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999

C:Accession: S48730

R:Namba, M.; Kurose, M.; Torioka, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.;

FEBS Lett. 353, 124-128, 1994

A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese ceda

A:Reference number: S48730; MUID:95010777; PMID:7926035

A:Accession: S48730

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-514 <KMW>

A:Cross-References: GB:D37765; NID:9577695; PIDN:BA07021.1; PID:dl007598; PID:957769

Query Match 100.0%; Score 67; DB 2; Length 514;

Best Local Similarity 100.0%; Pred. No. 0.00015;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSDISLKLTSKTIAS 15

Db 395 LSDISLKLTSKTIAS 409

RESULT 2

JC2498 second major allergen Cry j II precursor - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)

C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000

C:Accession: JC2498; PMID:8002972

R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994

A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese

A:Reference number: JC2498; MUID:94271186; PMID:8002972

A:Accession: JC2498

A:Molecule type: mRNA

A:Residues: 1-514 <KMW>

A:Cross-References: DDBJ:D29772; NID:9506857; PIDN:BA06172.1; PID:9506858

A:Accession: JC2498

A:Molecule type: protein

A:Residues: 52-61 <KO2>

R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matunashi, T.

Allergy 45, 309-312, 1990

A:Title: Identification of the second major allergen of Japanese cedar pollen.

A:Reference number: A60147; MUID:90342988; PMID:2382797

A:Accession: A60147

A:Molecule type: protein

A:Residues: 55-64 <SAK>

C:Keywords: glycoprotein; pollen

F:1-54/Domain: signal sequence

F:55-460/Product: second major allergen Cry j #status predicted <STG>

F:55-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 67; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSDISLKTSGKIAS 15
DB 395 LSDISLKTSGKIAS 409

RESULT 3

UC7100

polygalacturonase Cha o 2 - Japanese cypress

C/Species: Chamaecyparis obtusa (Japanese cypress)

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000

C/Accession: UC7100; PC7026

R/Motif: T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.

Biochem. Biophys. Res. Commun. 263, 166-171, 1999

A/Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all

A/Reference number: UC7100; MUID:19417540; PMID:10486272

A/Accession: UC7100

A/Molecule type: mRNA

A/Residues: 1-514 <MOR>

A/Accession: PC7026

A/Molecule type: protein

A/Residues: 51-62 <MO2>

Query Match

Best Local Similarity 80.6%; Score 54; DB 2; Length 514;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSDISLKTSGKIAS 15
DB 395 LSDISLKTSGKIAS 409

RESULT 4

UC7366

Jun a 2 protein - mountain cedar

C/Species: Juniperus ashei (mountain cedar)

C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000

C/Accession: UC7366; PC7093

R/Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.

Biochem. Biophys. Res. Commun. 275, 195-202, 2000

A/Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all

A/Reference number: UC7366

A/Accession: UC7366

A/Molecule type: mRNA

A/Residues: 1-507 <YOK>

A/Accession: PC7093

A/Molecule type: protein

A/Residues: 55-63 <YOK>

C/Comment: This protein, a second major allergen of mountain cedar pollen, which is inv

o the polygalacturonase family.

C/Keywords: glycoprotein; pollen

Query Match

Best Local Similarity 76.1%; Score 51; DB 2; Length 507;
Matches 10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSDISLKTSGKIAS 15
DB 396 LSDISLKTSGKIAS 410

RESULT 5

AH0742

high-affinity zinc uptake system ATP-binding protein [imported] - Salmonella enterica su

C/Species: Salmonella enterica subsp. enterica serovar Typhi

A/Note: This species has also been called Salmonella typhi

C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 17-May-2002

C/Accession: AH0742

R/Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Church

tn, T.; Comerford, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr

, S.; Mould, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K

A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ser

A/Reference number: AB0502; PMID:11677608

A/Accession: AH0742

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-251 <PAR>

A/Cross-references: GB:AL51382; PIDN:CAD05643.1; PID:G16503139; GSPDB:GN00176

C/Genetics:

A/Gene: STY2100

C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match

Best Local Similarity 67.2%; Score 45; DB 2; Length 251;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSDISLKTSGKI 13
DB 20 LSDVSLSPSKI 32

RESULT 6

P95256

ABC transporter, ATP-binding protein SP2196 [imported] - Streptococcus pneumoniae (str

C/Species: Streptococcus pneumoniae

C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C/Accession: P95256

R/Tetelid, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; He

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf

son, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris

A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A/Reference number: A95000; MUID:21357209; PMID:11463916

A/Accession: P95256

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-242 <KOR>

A/Cross-references: GB:AE005672; PIDN:AAK76247.1; PID:G14973708; GSPDB:GN00164; TIGR

A/Experimental source: strain TIGR4

C/Genetics:

A/Gene: SP2196

Query Match

Best Local Similarity 65.7%; Score 44; DB 2; Length 242;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSDISLKTSGKIAS 15
DB 19 LEDINLVTSSEVS 33

RESULT 7

G98121

hypothetical protein ABC-NBD [imported] - Streptococcus pneumoniae (strain R6)

C/Species: Streptococcus pneumoniae

C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001

C/Accession: G98121

R/Hockins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.

A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A/Reference number: A97872; MUID:21429245; PMID:11544234

A/Accession: G98121

A/Status: preliminary

A/Molecule type: DNA

A:Residues: 1-242 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAL00804.1; PID:g15459706; GSPDB:GN00174
 C:Genetics:
 A:Gene: ABC-NBD

Query Match 65.7%; Score 44; DB 2; Length 242;
 Best Local Similarity 53.3%; Pred. No. 1.5;
 Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSDISLKLTKSGKIAS 15
 DB 19 LEDIMLQVTSGEVVS 33

RESULT 8
 B64948
 Probable ABC transport system ATP-binding protein yebM - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:Accession: B64948
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 S:Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: B64948
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-251 <BLAT>
 A:Cross-references: GB:AE000280; GB:U00096; NID:g1788165; PIDN:AACT74928.1; PID:g1788165;
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: yebM
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 C:Keywords: ATP; nucleotide binding; P-loop; transport protein
 F:20-197/Domain: ATP-binding cassette homology <ABC>
 F:37-44/Region: nucleotide-binding motif A (P-loop)

Query Match 64.2%; Score 43; DB 2; Length 251;
 Best Local Similarity 69.2%; Pred. No. 2.5;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSDISLKLTKSGKI 13
 DB 20 LSDVSLKLKPKKI 32

RESULT 9
 D85798
 Hypothetical protein yebM [imported] - Escherichia coli (strain O157:H7, substrain EDL93
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
 C:Accession: D85798
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: D85798
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-251 <STO>
 A:Cross-references: GB:AE005174; NID:g12515914; PIDN:AAG56848.1; GSPDB:GN00145; UWGP:Z29
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: yebM
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 64.2%; Score 43; DB 2; Length 251;
 Best Local Similarity 69.2%; Pred. No. 2.5;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSDISLKLTKSGKI 13

DB 20 LSDVSLKLKPKKI 32

RESULT 10
 H90949
 Probable ABC transport system ATP-binding protein Ecs2568 [imported] - Escherichia co
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 27-Nov-2001
 C:Accession: H90949
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
 gasawara, N.; Yasunaga, T.; Kubara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
 A:Reference number: A99629; MUID:2116231; PMID:11258796
 A:Accession: H90949
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-251 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BAH35991.1; PID:g13362036; GSPDB:GN00154
 C:Experimental source: strain O157:H7, substrain RIMD 050952
 C:Genetics:
 A:Gene: Ecs2568

Query Match 64.2%; Score 43; DB 2; Length 251;
 Best Local Similarity 69.2%; Pred. No. 2.5;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSDISLKLTKSGKI 13
 DB 20 LSDVSLKLKPKKI 32

RESULT 11
 QRECM3
 Membrane-bound iron (III) dicitrate transport protein - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 01-Mar-2002
 C:Accession: J50115; S56512; A65242
 R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniele, D.L.; Blattner, F.R.
 Nucleic Acids Res. 23, 2105-2119, 1995
 A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from
 A:Reference number: S56512; MUID:95534362; PMID:7610040
 A:Accession: S56512
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-255 <BUR>
 A:Cross-references: EMBL:U14003; NID:g1263172; PIDN:AA97183.1; PID:g537128
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: A65242
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-255 <BLAT>
 A:Cross-references: GB:AE000499; GB:U00096; NID:g1790732; PIDN:AACT7243.1; PID:g179073
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: fecE
 A:Map position: 7 min
 C:Function:

A:Description: one of five, encoded by the fec operon, constituting a citrate-dependent C/Superfamily: inner membrane protein malK; ATP-binding cassette homology C/Keywords: ATP; iron transport; membrane protein; nucleotide binding; P-loop P:18-214/Domain: ATP-binding cassette homology <ABC> F:35-42/Region: nucleotide-binding motif A (P-loop)

Query Match 62.7%; Score 42; DB 1; Length 255;
Best Local Similarity 53.3%; Pred. No. 3.9;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSDISLKLTSKGIAS 15
DB 18 LNDVSLSLPTGKITA 32

RESULT 12

AI2834
Conserved hypothetical protein Atu2104 [imported] - Agrobacterium tumefaciens (strain C5
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C/Accession: AI2834

R/Mood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Grenchner, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A/Reference number: AB2577; PMID:11743193

A/Accession: AI2834

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-133 <KUR>

A/Cross-references: GB:AE008688; PIDN:AAI43095.1; PID:g17740566; GSPDB:GN00186

A/Experimental source: strain C58 (Dupont)

A/Genetics:

A/Map position: circular chromosome

Query Match 59.7%; Score 40; DB 2; Length 133;
Best Local Similarity 40.0%; Pred. No. 4.6;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSDISLKLTSKGIAS 15
DB 98 MTDLHLKITEGNVAA 112

RESULT 13

D97612
Hypothetical 14.2K protein in vacB-aidB intergenic region (0132) [imported] - Agrobacter
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C/Accession: D97612

R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A/Reference number: A97359; PMID:11743194

A/Accession: D97612

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-133 <KUR>

A/Cross-references: GB:AE007869; PIDN:AAK87653.1; PID:g15157239; GSPDB:GN00169

A/Genetics:

A/Map position: circular chromosome

Query Match 59.7%; Score 40; DB 2; Length 134;
Best Local Similarity 40.0%; Pred. No. 4.6;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSDISLKLTSKGIAS 15

DB 99 MTDLHLKITEGNVAA 113

RESULT 14

AD0251
High-affinity zinc uptake system ATP-binding protein [imported] - Yersinia pestis (str
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C/Accession: AD0251

R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; PMID:21470413; PMID:11586360

A/Accession: AD0251

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-253 <KUR>

A/Cross-references: GB:AL590842; PIDN:CAC90872.1; PID:g15980071; GSPDB:GN00175

A/Genetics:

A/Map position: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 59.7%; Score 40; DB 2; Length 253;
Best Local Similarity 69.2%; Pred. No. 9.3;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSDISLKLTSKGI 13
DB 21 LNDISLRLPGKI 33

RESULT 15

H90033
Hypothetical protein SA2132 [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: H90033

R/Kuroda, M.; Ohka, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogawara, N.; Hayashi, H.; Hiratake, K.
Lancet 357, 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A/Reference number: AB9758; PMID:21311952; PMID:11418146

A/Accession: H90033

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-299 <KUR>

A/Cross-references: GB:BA000018; PID:g13702141; PIDN:BAK43433.1; GSPDB:GN00149

A/Experimental source: strain N315

A/Genetics:

A/Map position: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 59.7%; Score 40; DB 2; Length 299;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSDISLKLTSKGI 13
DB 18 VNDISLRLSGKM 30

Search completed: April 20, 2003, 13:15:59
Job time: 8.07895 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds

(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524d-100

Sequence: 1 NNRIMQFAKLTGFT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A_Geneseq_101002.*

```
1: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	15	17	AA97887
2	82	100.0	105	18	AAW27370
3	82	100.0	134	18	AAW27371
4	82	100.0	460	16	AA69791
5	82	100.0	514	15	AA53690
6	82	100.0	514	16	AA74333
7	82	100.0	514	16	AA69792
8	82	100.0	514	17	AA93599
9	82	100.0	514	17	AA81586
10	82	100.0	514	20	AA25666

11	82	100.0	514	20	AA25667	Japanese cedar all
12	76	92.7	17	17	AA81581	Cedar pollen aller
13	70	85.4	17	17	AA81588	Cedar pollen aller
14	70	85.4	18	19	AAW80346	Sugi allergen prot
15	65	79.3	13	18	AAW12542	Japanese cedar pol
16	61	74.4	13	18	AAW12537	Japanese cedar pol
17	61	74.4	13	18	AAW12536	Japanese cedar pol
18	61	74.4	14	18	AAW12535	Japanese cedar pol
19	58	70.7	20	19	AAW42171	Japanese cedar pol
20	58	70.7	514	17	AAW04346	T-cell epitope pep
21	58	70.7	15	17	AAW42122	Chamaecyparis obtu
22	56	68.3	15	17	AAW97888	Japanese cypress p
23	56	68.3	15	19	AAW57761	Japanese cedar pollen
24	54	65.9	10	17	AAW81575	Residues 76-90 of
25	54	65.9	17	17	AAW81592	Cedar pollen aller
26	52	63.4	453	22	AAW51693	Juniperus ashei N-
27	52	63.4	507	22	AAW51691	Unipertus ashei Ju
28	50	61.0	11	18	AAW12543	Japanese cedar pol
29	50	61.0	12	18	AAW12544	Japanese cedar pol
30	50	61.0	15	17	AAW97888	Japanese cedar pollen
31	47	57.3	436	21	AAW52347	Arabidopsis thalia
32	47	57.3	459	21	AAW52346	Arabidopsis thalia
33	47	57.3	459	23	AAW91108	Herbicideally activ
34	46	56.1	10	18	AAW12545	Japanese cedar pol
35	46	56.1	11	18	AAW12546	Japanese cedar pol
36	46	56.1	12	18	AAW12547	Japanese cedar pol
37	46	56.1	12	22	AAW69107	Cryptomeria japoni
38	46	56.1	93	21	AAW3891	Artificial sequenc
39	46	56.1	93	21	AAW3896	Artificial sequenc
40	46	56.1	93	22	AAW69097	Cedar pollen aller
41	46	56.1	93	22	AAW69117	Cedar pollen aller
42	46	56.1	95	21	AAW3887	Artificial sequenc
43	46	56.1	95	21	AAW38901	Artificial sequenc
44	46	56.1	95	22	AAW69098	Cedar pollen aller
45	46	56.1	96	22	AAW69104	Cedar pollen aller

ALIGNMENTS

RESULT 1
AA97887 standard; peptide; 15 AA.
AA97887;
16-AUG-1996 (first entry)
Japan cedar pollen mature allergen Cry j II amino acids 81-95.
Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
Sugi pollinosis; diagnosis; treatment.
Cryptomeria japonica.
JP08047392-A.
20-FEB-1996.
07-NOV-1994; 94JP-0297840.
26-MAY-1994; 94JP-0134868.
05-NOV-1993; 93JP-0276773.
(MEIP) MEIJI MILK PROD CO LTD.
WPI, 1996-166249/17.
Japan cedar pollen allergen Cry j II epitope - comprises at least
part of specified 460 amino acid protein
Claim 8; Fig 3; 17pp; Japanese.

CC AAR97871-R97960 are overlapping peptides used for the epitope mapping
 CC of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic
 CC peptides of it are useful in the diagnosis, prevention and treatment
 CC of Sugi's pollinosis, the allergic reaction to Japan cedar pollen.
 CC Significant regions of the allergen were identified using the
 CC overlapping peptides of the full epitope derived from a Cry j II
 CC antigen-specific T cell line. Amino acids 66-80 (AAR97884) and 186-200
 CC (R978908) of the full mature 460 amino acid allergen are the most
 CC allergenic of the 90 peptides tested.

CC Sequence 15 AA;

Query Match 100.0%; Score 82; DB 17; Length 15;
 Best Local Similarity 100.0%; Pred. No. 9.7e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNRIMLOFAKLTGFT 15
 DB 1 NNRIMLOFAKLTGFT 15

RESULT 2
 AAM27370
 ID AAM27370 standard; peptide; 105 AA.

XX AAM27370;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #2.

XX Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 XX T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN WO9732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

PA (MEIP) MEIDI MILK PROD CO LTD.

PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;

XX WPI; 1997-470495/43.

PT Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens

PS Claim 6; Page 31; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as
 XX a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 XX or more different allergens (preferably linked via arginine or lysine
 XX dimers), where the T cell epitope regions have a positivity index
 XX greater than 100 as measured in a patient group responding to the
 XX allergen; have at least 70% reactivity with lymphocytes from patients
 XX responding to the allergen; and are not reactive with immunoglobulin E
 XX (IgE) antibodies from patients responsive to the allergen. The agent can
 XX be used to prevent and treat a wide variety of allergic diseases, e.g. by
 XX desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX Sequence 105 AA;

Query Match 100.0%; Score 82; DB 18; Length 105;
 Best Local Similarity 100.0%; Pred. No. 8.2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNRIMLOFAKLTGFT 15
 DB 54 NNRIMLOFAKLTGFT 68

RESULT 3
 AAM27371
 ID AAM27371 standard; peptide; 134 AA.

XX AAM27371;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #3.

XX Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 XX T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN WO9732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

PA (MEIP) MEIDI MILK PROD CO LTD.

PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;

XX WPI; 1997-470495/43.

PT Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens

PS Claim 6; Page 32; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as
 XX a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 XX or more different allergens (preferably linked via arginine or lysine
 XX dimers), where the T cell epitope regions have a positivity index
 XX greater than 100 as measured in a patient group responding to the
 XX allergen; have at least 70% reactivity with lymphocytes from patients
 XX responding to the allergen; and are not reactive with immunoglobulin E
 XX (IgE) antibodies from patients responsive to the allergen. The agent can
 XX be used to prevent and treat a wide variety of allergic diseases, e.g. by
 XX desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX Sequence 134 AA;

Query Match 100.0%; Score 82; DB 18; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNRIMLOFAKLTGFT 15
 DB 54 NNRIMLOFAKLTGFT 68

RESULT 4

AAR69791
 ID AAR69791 standard; Protein; 460 AA.

XX AAR69791;

DT 27-SEP-1995 (first entry)

DE Japonicum allergen residues 55-514.

XX Japonicum allergen; residues 55-514; induced histamine release;

KW antiallergic peptide; IGE cross-linking inhibition.
XX Japonicum sp.
OS
XX WO502412-A.
XX
XX 26-JAN-1995.
XX
XX 15-JUL-1994; 94WO-JP01164.
XX
XX 16-JUL-1993; 93JP-0177008.
XX 01-SEP-1993; 93JP-0217725.
XX 07-APR-1994; 94JP-0069336.
XX
XX (MEIP) MEIJI MILK PROD CO LTD.
XX
XX Kino K, Kohno Y, Komiyama N, Sone T;
XX
XX WPI; 1995-067159/09.
XX N-PSDB; AA084044.
XX
XX Peptide antiallergic agent - inhibits cross-linking of allergen
XX with IGE antibody
XX
XX Disclosure; Pages 26-27; 46pp; Japanese.
XX
XX AA084044 encodes AAR69791 Japonicum allergen residues 55-514, from
XX which the antiallergic peptides AAR69845-R69809 were derived.
XX The peptides ability to inhibit the cross-linking of an allergen,
XX to an IGE antibody can be used in the prevention and treatment of
XX allergic diseases.
XX
XX Sequence 460 AA;
SQ
Query Match 100.0%; Score 82; DB 16; Length 460;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NNRIMLOFAKLTGFT 15
DB 81 NNRIMLOFAKLTGFT 95
RESULT 5
AAR53690
ID AAR53690 standard; Protein; 514 AA.
XX
XX AAR53690;
AC
XX 01-FEB-1995 (first entry)
XX
XX Japanese cedar pollen allergen Cry j II.
XX
XX Cedar pollinosis; diagnostic.
XX
XX Cryptomeria japonica.
XX
XX WO9411512-A.
XX
XX 26-MAY-1994.
XX
XX 12-NOV-1993; 93WO-US11000.
XX
XX 12-NOV-1992; 92US-0975179.
XX
XX (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
XX Brauer A, Kuo M, Pollock J, Yeung S;
XX WPI; 1994-183513/22.
XX N-PSDB; AA066048.
XX
XX Allergenic Cry j II protein and fragments from Japanese cedar

PT pollen - used to diagnose, treat and prevent Japanese cedar
PT pollinosis
XX Claim 2; Fig 4; 89pp; English.
XX
XX The sequence is of a Japanese cedar pollen allergen Cry j
XX II. The protein and its fragments can be used for diagnosis and
XX treatment of Japanese cedar pollinosis and to identify similar
XX sequences in other plants.
XX See also AAR53692-6.
XX
XX Sequence 514 AA;
SQ
Query Match 100.0%; Score 82; DB 15; Length 514;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NNRIMLOFAKLTGFT 15
DB 135 NNRIMLOFAKLTGFT 149
RESULT 6
AAR74333
ID AAR74333 standard; Protein; 514 AA.
XX
XX AAR74333;
AC
XX 01-NOV-1995 (first entry)
XX
XX Japanese cedar pollen allergen.
XX
XX Japanese cedar; pollen; allergen; allergy; therapy; diagnostic;
XX desensitizer.
XX
XX Cryptomeria japonica.
XX
XX EP655500-A.
XX
XX 31-MAY-1995.
XX
XX 03-NOV-1994; 94EP-0308117.
XX
XX 27-DEC-1993; 93JP-0346814.
XX 05-NOV-1993; 93JP-0299151.
XX 20-DEC-1993; 93JP-0344596.
XX
XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Kurimoto M, Namba M, Torigoe K;
XX
XX WPI; 1995-195588/26.
XX N-PSDB; AA090156.
XX
XX New Japanese cedar pollen allergen polypeptide - and DNA coding for it,
XX useful for treatment and diagnosis of cedar pollen allergy
XX
XX Claim 5; Page 26-28; 41pp; English.
XX
XX The gene encoding an allergen of Japanese cedar pollen was isolated
XX by PCR amplification using primers based on portions of the allergen
XX protein. The gene was used for recombinant allergen production in
XX E. coli (vector plasmid pKK-223-3).
XX
XX Sequence 514 AA;
SQ
Query Match 100.0%; Score 82; DB 16; Length 514;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NNRIMLOFAKLTGFT 15
DB 135 NNRIMLOFAKLTGFT 149

FT	Protein	55..514	
FT		/*label= mat_protein	
XX			
XX	JP08047392-A.		
XX			
XX	20-FEB-1996.		
XX			
XX	07-NOV-1994;	94JP-0297840.	
XX			
XX	26-MAY-1994;	94JP-0134868.	
XX	05-NOV-1993;	93JP-0276773.	
XX			
XX	(MEIP) MEIJI MILK PROD CO LTD.		
XX			
XX	WPI; 1996-166249/17.		
XX	N-PSDB; AAT18102.		
XX			
XX	Japan cedar pollen allergen Cry j II epitope - comprises at least		
XX	part of specified 460 amino acid protein		
XX			
XX	Claim 1; Page 10-11; 17pp; Japanese.		
XX			
XX	AA893599 is a Japan cedar pollen Cry j II allergen which is useful		
XX	in the diagnosis, prevention and treatment of Sugi pollinosis,		
XX	the allergic reaction to Japan cedar pollen. Significant regions of		
XX	the allergen were identified using overlapping peptides of the full		
XX	epitope derived from a Cry j II antigen-specific T cell line		
XX	(see AA897871-R97960). Amino acids 68-80 (AA897884) and 186-200		
XX	(R9789908) of the full mature 460 amino acid allergen are the most		
XX	allergenic of the 90 peptides tested.		
XX			
XX	Sequence 514 AA;		
XX			
XX	Query Match	100.0%; Score 82; DB 17; Length 514;	
XX	Best Local Similarity	100.0%; Pred. No. 4.7e-06;	
XX	Matches 15; Conservative	0; Mismatches 0; Indels 0; Gaps	0;
XX			
XX	1 NNRITWQFAKLTGFT 15		
XX	135 NNRITWQFAKLTGFT 149		
XX			
XX			
XX	RESULT 9		
XX	AA81586		
XX	ID AA81586 standard; Protein; 514 AA.		
XX			
XX	AA81586;		
XX			
XX	24-MAY-1996 (first entry)		
XX			
XX	Cedar pollen allergen A.		
XX			
XX	Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;		
XX	KW antibody; pollinosis; therapy; immunotherapy.		
XX			
XX	Cryptomeria japonica.		
XX			
XX	EP700929-A2.		
XX			
XX	13-MAR-1996.		
XX			
XX	08-SEP-1995; 95EP-0306295.		
XX			
XX	14-JUL-1995; 95JP-0200221.		
XX	10-SEP-1994; 94JP-0242137.		
XX	14-JUL-1995; 95JP-0200204.		
XX			
XX	(HAYB) HAYASHIBARA SEIBUTSU KAGAKU.		
XX			
XX	Hino K, Saito S, Taniguchi Y;		
XX			
XX	WPI; 1996-140976/15.		
XX			

PT New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IGE antibodies,
 used for treating cedar pollinosis
 XX
 PS Claim 5; Page 29-30; 36pp; English.
 CC
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (AAR81586) and B (AAR81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IGE
 CC antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
 CC for T-cell recognition, and homologous peptides (AAR81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 XX
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 82; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 4,7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NNRIMLOFAKLTGFT 15
 DB 135 NNRIMLOFAKLTGFT 149
 RESULT 10
 AAY25666
 XX AAY25666 standard; protein; 514 AA.
 AC
 XX AAY25666;
 DT 30-SEP-1999 (first entry)
 DE Japanese cedar allergen 1076242 Cry j II precursor protein fragment.
 XX
 KM Major histocompatibility complex; class II; desensitizing; human;
 KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX
 OS Cedrus sp.
 XX
 PN WO934826-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 11-JAN-1999; 99WO-GB00080.
 XX
 XX 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Kay AB, Larche M;
 XX
 DR WPI; 1999-458255/38.
 XX
 PT Desensitizing patients to polypeptide allergens
 XX
 PS Example 6; Page 74; 117pp; English.
 CC
 CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit

CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076242 Cry j II
 CC precursor.
 XX
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 82; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 4,7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NNRIMLOFAKLTGFT 15
 DB 135 NNRIMLOFAKLTGFT 149
 RESULT 11
 AAY25667
 XX AAY25667 standard; protein; 514 AA.
 AC
 XX AAY25667;
 DT 30-SEP-1999 (first entry)
 DE Japanese cedar allergen 1076241 Cry j II protein fragment.
 XX
 KM Major histocompatibility complex; class II; desensitizing; human;
 KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX
 OS Cedrus sp.
 XX
 PN WO934826-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 11-JAN-1999; 99WO-GB00080.
 XX
 XX 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Kay AB, Larche M;
 XX
 DR WPI; 1999-458255/38.
 XX
 PT Desensitizing patients to polypeptide allergens
 XX
 PS Example 6; Page 74; 117pp; English.
 CC
 CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence

CC represents a Japanese cedar (Cedrus sp.) allergen 1076241 Cry j II.

XX
SQ Sequence 514 AA;

Query Match 100.0%; Score 82; DB 20; Length 514;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NNRIMLOFAKLTGFT 15
DB 135 NNRIMLOFAKLTGFT 149

RESULT 12
AAR81581
ID AAR81581 standard; Peptide, 17 AA.

XX
AC AAR81581;
XX
AC AAR81581;
XX
DT 24-MAY-1996 (first entry)
XX
DE Cedar pollen allergen peptide 9 (T-cell epitope).
XX
DE Cedar; Cryptomeria japonica; pollen; allergen; immunoglobulin E;
XX
KW IGE; T-cell epitope; antibody; pollinosis; therapy; immunotherapy.
XX
OS Synthetic.
XX
PN EP700929-A2.
XX
PD 13-MAR-1996.
XX
PP 08-SEP-1995; 95EP-0306295.
XX
PR 14-JUL-1995; 95JP-0200221.
XX
PR 10-SEP-1994; 94JP-0242137.
XX
PR 14-JUL-1995; 95JP-0200204.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hino K, Saito S, Taniguchi Y;
XX
DR WPI; 1996-140976/15.
XX
PT New peptide(s) derived from cedar pollen allergens - activate
XX
PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
XX
PT used for treating cedar pollinosis
XX
PS Claim 4; Page 28; 36pp; English.
XX
CC Synthetic peptides based on portions of cedar pollen allergens A
XX
CC (AAR81586) and B (AAR81587) were tested for their ability to activate
XX
CC cedar allergen-specific T-cells, but not allergen-specific IgE
XX
CC antibodies. 6 Peptides (AAR81580-R81585) were identified as T-cell
XX
CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
XX
CC for T-cell recognition, and homologous peptides (AAR81588-96) can
XX
CC be used as immunotherapeutic agents to treat or prevent cedar
XX
CC pollinosis, avoiding side-effects such as anaphylaxis.
XX
SQ Sequence 17 AA;

Query Match 92.7%; Score 76; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 NNRIMLOFAKLTGFT 15
DB 1 NNRIMLOFAKLTGFT 14

RESULT 13
AAR81588
ID AAR81588 standard; Peptide, 17 AA.

XX
AC AAR81588;
XX
DT 24-MAY-1996 (first entry)
XX
DE Cedar pollen allergen peptide 16 (homologue).
XX
DE Cedar; Cryptomeria japonica; pollen; allergen; immunoglobulin E;
XX
KW IGE; T-cell epitope; antibody; pollinosis; therapy; immunotherapy.
XX
OS Synthetic.
XX
PN EP700929-A2.
XX
PD 13-MAR-1996.
XX
PP 08-SEP-1995; 95EP-0306295.
XX
PR 14-JUL-1995; 95JP-0200221.
XX
PR 10-SEP-1994; 94JP-0242137.
XX
PR 14-JUL-1995; 95JP-0200204.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hino K, Saito S, Taniguchi Y;
XX
DR WPI; 1996-140976/15.
XX
PT New peptide(s) derived from cedar pollen allergens - activate
XX
PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
XX
PT used for treating cedar pollinosis
XX
PS Disclosure; Page 32; 36pp; English.
XX
CC Synthetic peptides based on portions of cedar pollen allergens A
XX
CC (AAR81586) and B (AAR81587) were tested for their ability to activate
XX
CC cedar allergen-specific T-cells, but not allergen-specific IgE
XX
CC antibodies. 6 Peptides (AAR81580-R81585) were identified as T-cell
XX
CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
XX
CC for T-cell recognition, and homologous peptides (AAR81588-96)
XX
CC that contain 1 or more amino acid substitution(s) can be used
XX
CC as immunotherapeutic agents to treat or prevent cedar pollinosis,
XX
CC avoiding side-effects such as anaphylaxis.
XX
SQ Sequence 17 AA;

Query Match 85.4%; Score 70; DB 17; Length 17;
Best Local Similarity 92.9%; Pred. No. 1.4e-05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 NNRIMLOFAKLTGFT 15
DB 1 NNRIMLOFAKLTGFT 14

RESULT 14
AAR80346
ID AAR80346 standard; peptide, 18 AA.

XX
AC AAR80346;
XX
DT 11-JAN-1999 (first entry)
XX
DE Sugi allergen protein Cryj2 derived epitope for T cells.
XX
DE T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;
XX
KW sugi-pollinosis; allergic reaction; pollen.
XX
OS Synthetic.
XX
PN JP10259198-A.
XX
PD 29-SEP-1998.

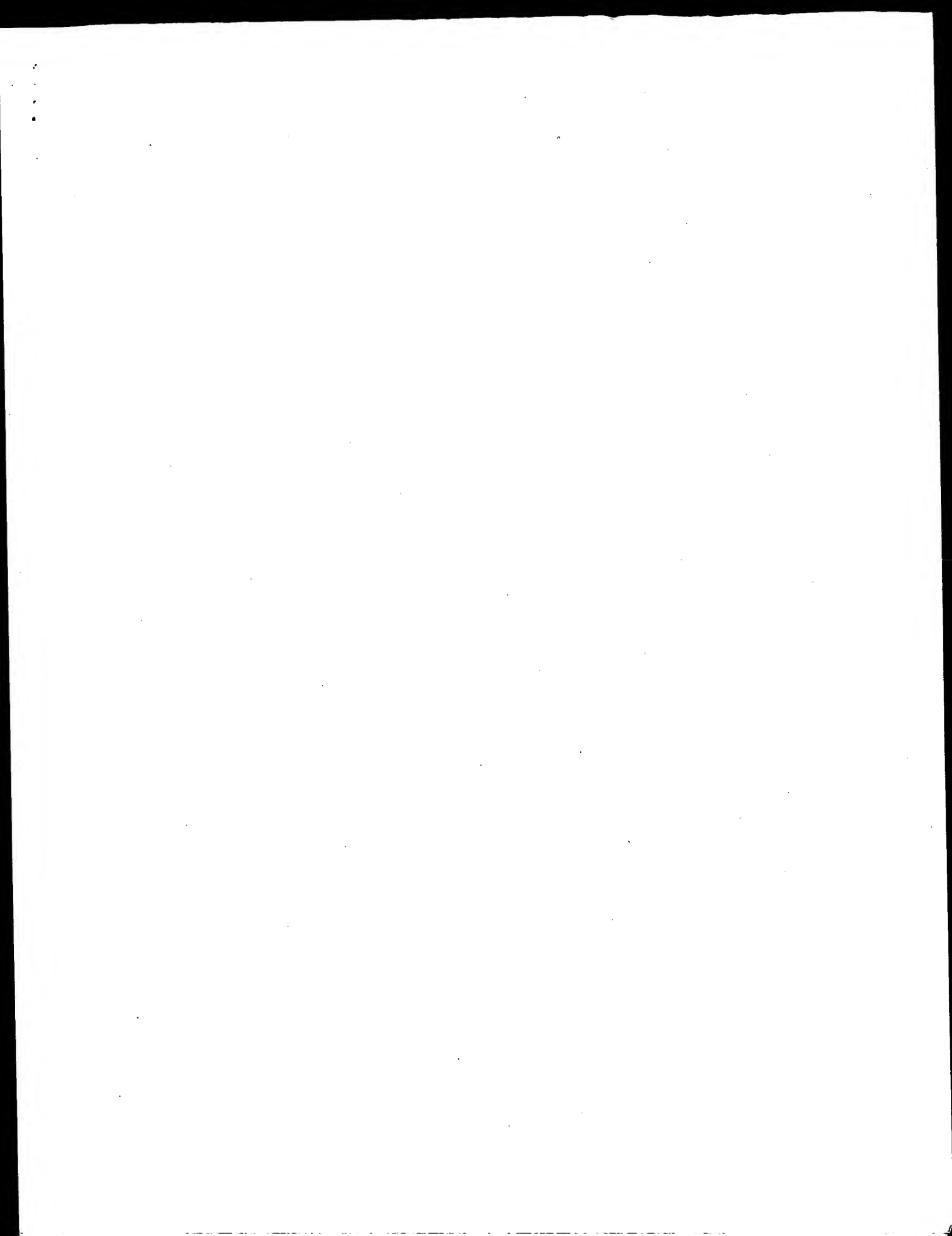
XX 22-DEC-1997; 97JP-0353448.
 PF 24-DEC-1996; 96JP-0343441.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX (SANY) SANKYO CO LTD.
 DR WPI; 1998-577037/49.
 XX
 PT A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollinosis
 XX
 PS Claim 7; Page 18; 21pp; Japanese.
 CC
 CC AAW80339-58 represent epitopes for T cells, derived from the sugi
 CC allergen proteins Cryj1 (AAW80339-44, AAW80350-53 and AAW80356-58) and
 CC Cryj2 (AAW80345-49 and AAW80354-55). The peptides are useful for the
 CC treatment of sugi-pollinosis, an allergic reaction of the body to
 CC pollen.
 CC
 SQ Sequence 18 AA;

Query Match 85.4%; Score 70; DB 19; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RIWLOFAKLTGFT 15
 Db 1 RIWLOFAKLTGFT 13

RESULT 15
 AAW12542
 ID AAW12542 standard; peptide; 13 AA.
 AC AAW12542;
 XX
 DT 30-APR-1997 (first entry)
 XX
 DE Japanese cedar pollen allergen-derived peptide 11.
 XX
 KW Japanese cedar; sugi; pollen; pollinosis; pollenosis; allergen;
 KW allergy; T-cell epitope; TCR; T-cell receptor; activation;
 KW immune tolerance; Cryj.
 XX
 OS Cryptomeria japonica.
 OS Synthetic.
 XX
 PN JP0833391-A.
 PD 17-DEC-1996.
 XX
 PF 18-JUL-1995; 95JP-0181438.
 XX
 PR 07-APR-1995; 95JP-0082519.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 DR WPI; 1997-095487/09.
 XX
 PT Peptide allergen derived from Japanese cedar pollen - causes T cell
 PT response specific to cedar pollen, for treatment of pollenosis
 XX
 PS Claim 16; Page 19; 21pp; Japanese.
 CC
 CC The present sequence is one of 24 claimed peptides which were
 CC synthesised based on Japanese cedar pollen sequences. This peptide
 CC was shown to have Japanese cedar pollen antigen T cell epitope
 CC activity by using T cells isolated from a cedar pollinosis patient.
 CC The peptide produces little or no anaphylaxis.
 CC

SQ Sequence 13 AA;
 Query Match 79.3%; Score 65; DB 18; Length 13;
 Best Local Similarity 100.0%; Pred. No. 8.2e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 IWLOFAKLTGFT 15
 Db 1 IWLOFAKLTGFT 12

Search completed: April 20, 2003, 13:06:09
 Job time : 18.1974 secs



OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds

(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524D-101

Sequence: 1 LQFAKL TGF TLMGKG 15

Scoring table: BLOSUM62

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
```

Database : A_Geneseq_101002: *

1	/SID2_gcgdata/geneseq/geneexp-emb1/AA1980.DAT.*
2	/SID2_gcgdata/geneseq/geneexp-emb1/AA1981.DAT.*
3	/SID2_gcgdata/geneseq/geneexp-emb1/AA1982.DAT.*
4	/SID2_gcgdata/geneseq/geneexp-emb1/AA1983.DAT.*
5	/SID2_gcgdata/geneseq/geneexp-emb1/AA1984.DAT.*
6	/SID2_gcgdata/geneseq/geneexp-emb1/AA1985.DAT.*
7	/SID2_gcgdata/geneseq/geneexp-emb1/AA1986.DAT.*
8	/SID2_gcgdata/geneseq/geneexp-emb1/AA1987.DAT.*
9	/SID2_gcgdata/geneseq/geneexp-emb1/AA1988.DAT.*
10	/SID2_gcgdata/geneseq/geneexp-emb1/AA1990.DAT.*
11	/SID2_gcgdata/geneseq/geneexp-emb1/AA1990.DAT.*
12	/SID2_gcgdata/geneseq/geneexp-emb1/AA1991.DAT.*
13	/SID2_gcgdata/geneseq/geneexp-emb1/AA1992.DAT.*
14	/SID2_gcgdata/geneseq/geneexp-emb1/AA1993.DAT.*
15	/SID2_gcgdata/geneseq/geneexp-emb1/AA1994.DAT.*
16	/SID2_gcgdata/geneseq/geneexp-emb1/AA1995.DAT.*
17	/SID2_gcgdata/geneseq/geneexp-emb1/AA1996.DAT.*
18	/SID2_gcgdata/geneseq/geneexp-emb1/AA1997.DAT.*
19	/SID2_gcgdata/geneseq/geneexp-emb1/AA1998.DAT.*
20	/SID2_gcgdata/geneseq/geneexp-emb1/AA1999.DAT.*
21	/SID2_gcgdata/geneseq/geneexp-emb1/AA2000.DAT.*
22	/SID2_gcgdata/geneseq/geneexp-emb1/AA2001.DAT.*
23	/SID2_gcgdata/geneseq/geneexp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	76	100.0	15	17	AAR97888	Japan cedar pollen
2	76	100.0	18	19	AAW80346	Sugi allergen prot
3	76	100.0	460	16	AAR69791	Japonicum allergen
4	76	100.0	514	15	AAR53690	Japanese cedar pol
5	76	100.0	514	16	AAR74333	Japanese cedar pol
6	76	100.0	514	16	AAR69792	Japonicum allergen
7	76	100.0	514	17	AAR93599	Japan cedar pollen
8	76	100.0	514	20	AAR81586	Cedar pollen allerg
9	76	100.0	514	17	AAV25666	Japanese cedar all
10	76	100.0	514	20	AAV25667	Japanese cedar all

11	67	88.2	105	18	AAW27370	Multi-epitope pepb
12	67	88.2	134	18	AAW27371	Multi-epitope pepb
13	65	85.5	14	18	AAW12535	Japanese cedar pol
14	65	85.5	17	17	AAW81581	Cedar pollen aller
15	63	82.9	16	22	AAW69104	Cedar pollen aller
16	63	82.9	96	22	AAW69119	Cedar pollen aller
17	63	82.9	97	22	AAW69105	Cedar pollen aller
18	63	82.9	210	22	AAW69103	Cedar pollen aller
19	63	82.9	214	22	AAW69120	Cedar pollen aller
20	61	80.3	12	18	AAW12547	Japanese cedar pol
21	61	80.3	12	22	AAW69107	Cryptomeria japoni
22	61	80.3	93	21	AAW23891	Artificial sequenc
23	61	80.3	93	21	AAW23896	Artificial sequenc
24	61	80.3	93	22	AAW69097	Cedar pollen aller
25	61	80.3	93	22	AAW69117	Cedar pollen aller
26	61	80.3	95	21	AAW23897	Artificial sequenc
27	61	80.3	95	21	AAW23901	Artificial sequenc
28	61	80.3	95	22	AAW69098	Cedar pollen aller
29	61	80.3	99	21	AAW23802	Artificial sequenc
30	61	80.3	99	21	AAW23907	Cedar pollen aller
31	61	80.3	99	22	AAW69099	Artificial sequenc
32	59	77.6	12	18	AAW12544	Cedar pollen aller
33	59	77.6	13	18	AAW12536	Japanese cedar pol
34	59	77.6	17	17	AAW81588	Japanese cedar pol
35	56	73.7	94	21	AAW23908	Cedar pollen aller
36	55	72.4	11	18	AAW12546	Artificial sequenc
37	54	71.1	11	18	AAW12543	Japanese cedar pol
38	54	71.1	12	18	AAW12537	Japanese cedar pol
39	54	71.1	13	18	AAW12542	Japanese cedar pol
40	54	71.1	17	17	AAW04346	Japanese cedar pol
41	54	71.1	514	19	AAW42122	Chamaecyparis obtu
42	52	68.4	15	17	AAW97889	Japanese cypress p
43	50	65.8	10	18	AAW12545	Japanese cedar pol
44	50	65.8	15	17	AAW97887	Japanese cedar pol
45	50	65.8	20	19	AAW42172	T-cell epitope pep

ALIGNMENTS

RESULT 1
AAR97888

DT 16-AUG-1996 (first entry)
 XY

Japan cedar pollen mature allergen Cry j II amino acids 86-100.

KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen.
 KW Sugi pollinosis; diagnosis; treatment.
 KW

OS Cryptomeria japonica.
XX

PN JP08047392-A
XX

PD 20-FEB-1996
XX

PF 07~NOV-1994; 94JP-0297840
XX

PR	26-MAY-1994;	94JP-0134868
PR	05-NOV-1993;	93JP-0376773

XX
PA
(METP) MET.IT MIK PROD CO LTD

XX
DR WPT: 1996-166249/17

xx Japan cedar pollen allergy symptoms - common at 1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-20-21-22-23-24-25-26-27-28-29-30-31-32-33-34-35-36-37-38-39-40-41-42-43-44-45-46-47-48-49-50-51-52-53-54-55-56-57-58-59-60-61-62-63-64-65-66-67-68-69-70-71-72-73-74-75-76-77-78-79-80-81-82-83-84-85-86-87-88-89-90-91-92-93-94-95-96-97-98-99-100-101-102-103-104-105-106-107-108-109-110-111-112-113-114-115-116-117-118-119-120-121-122-123-124-125-126-127-128-129-130-131-132-133-134-135-136-137-138-139-140-141-142-143-144-145-146-147-148-149-150-151-152-153-154-155-156-157-158-159-160-161-162-163-164-165-166-167-168-169-170-171-172-173-174-175-176-177-178-179-180-181-182-183-184-185-186-187-188-189-190-191-192-193-194-195-196-197-198-199-200-201-202-203-204-205-206-207-208-209-210-211-212-213-214-215-216-217-218-219-220-221-222-223-224-225-226-227-228-229-230-231-232-233-234-235-236-237-238-239-240-241-242-243-244-245-246-247-248-249-250-251-252-253-254-255-256-257-258-259-260-261-262-263-264-265-266-267-268-269-270-271-272-273-274-275-276-277-278-279-280-281-282-283-284-285-286-287-288-289-290-291-292-293-294-295-296-297-298-299-300-301-302-303-304-305-306-307-308-309-310-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-896-897-898-899-900-901-902-903-904-905-906-907-908-909-910-911-912-913-914-915-916-917-918-919-920-921-922-923-924-925-926-927-928-929-930-931-932-933-934-935-936-937-938-939-940-941-942-943-944-945-946-947-948-949-950-951-952-953-954-955-956-957-958-959-960-961-962-963-964-965-966-967-968-969-970-971-972-973-974-975-976-977-978-979-980-981-982-983-984-985-986-987-988-989-990-991-992-993-994-995-996-997-998-999-1000-1001-1002-1003-1004-1005-1006-1007-1008-1009-1010-1011-1012-1013-1014-1015-1016-1017-1018-1019-1020-1021-1022-1023-1024-1025-1026-1027-1028-1029-1030-1031-1032-1033-1034-1035-1036-1037-1038-1

part of specified 460 amino acid protein

claim of 1/1 pp; Japanese.
XX

AA697871-R97960 are overlapping peptides used for the epitope mapping of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic peptides of it are useful in the diagnosis, prevention and treatment of Sugi pollinosis, the allergic reaction to Japan cedar pollen. Significant regions of the allergen were identified using the overlapping peptides of the full epitope derived from a Cry j II antigen-specific T cell line. Amino acids 66-80 (AA697884) and 186-200 (R978908) of the full mature 460 amino acid allergen are the most allergenic of the 90 peptides tested.

Sequence 15 AA;

Query Match 100.0%; Score 76; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOPAKLTGFTLMGKG 15
1 LOPAKLTGFTLMGKG 15

RESULT 2
AA680346
ID AA680346 standard; peptide; 18 AA.

AC AA680346;

DT 11-JAN-1999 (first entry)

DE Sugi allergen protein Cryj2 derived epitope for T cells.

KW T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment; sugi-pollinosis; allergic reaction; pollen.

OS Synthetic.

PN JP10259198-A.

PD 29-SEP-1998.

PF 22-DEC-1997; 97JP-0353448.

PR 24-DEC-1996; 96JP-0343441.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

PA (SANY) SANKYO CO LTD.

DR WPI; 1998-577037/49.

PT A linked T cell epitope peptide - used for the treatment of sugi-pollinosis

PS Claim 7; Page 18; 21pp; Japanese.

AA680339-58 represent epitopes for T cells, derived from the sugi allergen proteins Cryj1 (AA680339-44, AA680350-53 and AA680356-58) and Cryj2 (AA680345-49 and AA680354-55). The peptides are useful for the treatment of sugi-pollinosis, an allergic reaction of the body to pollen.

Sequence 18 AA;

Query Match 100.0%; Score 76; DB 19; Length 18;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOPAKLTGFTLMGKG 15
4 LOPAKLTGFTLMGKG 18

RESULT 3
AA69791

ID AA69791 standard; Protein; 460 AA.

AC AA69791;

DT 27-SEP-1995 (first entry)

DE Japonicum allergen residues 55-514.

KW Japonicum allergen; residues 55-514; induced histamine release; anti-allergic peptide; IgE cross-linking inhibition.

OS Japonicum sp.

PN WO9502412-A.

PD 26-JAN-1995.

PF 15-JUL-1994; 94WO-JP01164.

PR 16-JUL-1993; 93JP-0177008.

PR 01-SEP-1993; 93JP-0217725.

PR 07-APR-1994; 94JP-0069336.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Kuno K, Kohno Y, Komiyama N, Sone T;

DR WPI; 1995-067159/09.

DR N-PSDB; AA084044.

PT Peptide anti-allergic agent - inhibits cross-linking of allergen with IgE antibody

PS Disclosure; Pages 26-27; 46pp; Japanese.

AA084044 encodes AA69791 Japonicum allergen residues 55-514, from which the anti-allergic peptides AA69845-R69809 were derived. The peptides ability to inhibit the cross-linking of an allergen, to an IgE antibody can be used in the prevention and treatment of allergic diseases.

Sequence 460 AA;

Query Match 100.0%; Score 76; DB 16; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOPAKLTGFTLMGKG 15
86 LOPAKLTGFTLMGKG 100

RESULT 4
AA653690

ID AA653690 standard; Protein; 514 AA.

AC AA653690;

DT 01-FEB-1995 (first entry)

DE Japanese cedar pollen allergen Cry j II.

KW Cedar pollinosis; diagnostic.

OS Cryptomeria japonica.

PN WO9411512-A.

PD 26-MAY-1994.

PF 12-NOV-1993; 93WO-US11000.

PR 12-NOV-1992; 92US-0975179.

XX (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX Brauer A, Kuo M, Pollock J, Yeung S;
 XX WPI; 1994-183513/22.
 DR N-PSDB; AAQ66048.
 XX
 XX Allergenic Cry j II protein and fragments from Japanese cedar
 PT pollen - used to diagnose, treat and prevent Japanese cedar
 PT pollinosis
 XX
 PS Claim 2; Fig 4; 89pp; English.
 CC The sequence is of a Japanese cedar pollen allergen Cry j
 CC II. The protein and its fragments can be used for diagnosis and
 CC treatment of Japanese cedar pollinosis and to identify similar
 CC sequences in other plants.
 CC See also AAR53692-6.
 CC
 SQ Sequence 514 AA;

Query Match 100.0%; Score 76; DB 15; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LOPAKLTGFTLMGKG 15
 DB 140 LOPAKLTGFTLMGKG 154

RESULT 5
 AAR74333
 ID AAR74333 standard; Protein; 514 AA.
 XX
 AC AAR74333;
 XX
 DT 01-NOV-1995 (first entry)
 XX
 DE Japanese cedar pollen allergen.
 XX
 KM Japanese cedar; pollen; allergen; allergy; therapy; diagnostic;
 KM desensitizer.
 XX
 OS Cryptomeria japonica.
 XX
 PN EP655500-A.
 XX
 PD 31-MAY-1995.
 XX
 PF 03-NOV-1994; 94EP-0308117.
 XX
 PR 27-DEC-1993; 93JP-0346814.
 PR 05-NOV-1993; 93JP-0299151.
 PR 20-DEC-1993; 93JP-0344596.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 PI Kurimoto M, Namba M, Torigoe K;
 XX
 DR WPI; 1995-195588/26.
 DR N-PSDB; AAQ90156.
 XX
 PT New Japanese cedar pollen allergen polypeptide - and DNA coding for it,
 PT useful for treatment and diagnosis of cedar pollen allergy
 XX
 PS Claim 5; Page 26-28; 41pp; English.
 CC The gene encoding an allergen of Japanese cedar pollen was isolated
 CC by PCR amplification using primers based on portions of the allergen
 CC protein. The gene was used for recombinant allergen production in
 CC E. coli (vector plasmid pKK-223-3).
 XX

SQ Sequence 514 AA;
 Query Match 100.0%; Score 76; DB 16; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LOPAKLTGFTLMGKG 15
 DB 140 LOPAKLTGFTLMGKG 154

RESULT 6
 AAR69792
 ID AAR69792 standard; Protein; 514 AA.
 XX
 AC AAR69792;
 XX
 DT 27-SEP-1995 (first entry)
 XX
 DE Japonicum allergen.
 XX
 KM Japonicum allergen; induced histamine release; antiallergic peptide;
 KM IGE cross-linking inhibition.
 XX
 OS Japonicum sp.
 XX
 PN WO9502412-A.
 XX
 PD 26-JAN-1995.
 XX
 PF 15-JUL-1994; 94WO-JP01164.
 XX
 PR 16-JUL-1993; 93JP-0177008.
 PR 01-SEP-1993; 93JP-0217725.
 PR 07-APR-1994; 94JP-0069336.
 XX
 PA (MEIP) MEIJI MILK PROD CO LTD.
 XX
 PI Kino K, Kohno Y, Komiyama N, Sone T;
 XX
 DR WPI; 1995-067159/09.
 DR N-PSDB; AAQ84045; AAQ84046.
 XX

PT Peptide antiallergic agent - inhibits cross-linking of allergen
 PT with IGE antibody
 XX
 PS Example 3; Pages 27-28; 46pp; Japanese.
 XX
 CC AAQ84045 encodes AAR69792 Japonicum allergen, from which the
 CC antiallergic peptides AAR69845-R69809 were derived. The peptides ability
 CC to inhibit the cross-linking of an allergen, to an IGE antibody can be
 CC used in the prevention and treatment of allergic diseases.
 XX
 SQ Sequence 514 AA;

Query Match 100.0%; Score 76; DB 16; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LOPAKLTGFTLMGKG 15
 DB 140 LOPAKLTGFTLMGKG 154

RESULT 7
 AAR93599
 ID AAR93599 standard; Protein; 514 AA.
 XX
 AC AAR93599;
 XX
 DT 16-AUG-1996 (first entry)
 XX
 DE Japan cedar pollen Cry j II allergen.

XX Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
 KW Sugi pollinosis; diagnosis; treatment.
 XX Cryptomeria japonica.
 XX Key Location/Qualifiers
 FT Peptide 1..54
 FT /*label= sig_peptide
 FT Protein 55..514
 FT /*label= mat_protein
 XX JP08047392-A.
 PD 20-FEB-1996.
 XX 07-NOV-1994; 94UP-0297840.
 XX 26-MAY-1994; 94UP-0134868.
 XX 05-NOV-1993; 93UP-0276773.
 XX (MEIP) MEIJI MILK PROD CO LTD.
 XX WPI; 1996-166249/17.
 XX N-PSDB; AAT18102.
 XX Japan cedar pollen allergen Cry j II epitope - comprises at least
 PT part of specified 460 amino acid protein
 XX Claim 1; Page 10-11; 17pp; Japanese.
 XX AAR93599 is a Japan cedar pollen Cry j II allergen which is useful
 CC in the diagnosis, prevention and treatment of Sugi pollinosis,
 CC the allergic reaction to Japan cedar pollen. Significant regions of
 CC the allergen were identified using overlapping peptides of the full
 CC epitope derived from a Cry j II antigen-specific T cell line
 CC (see AAR97871-R97960). Amino acids 66-80 (AAR97884) and 186-200
 CC (R978908) of the full mature 460 amino acid allergen are the most
 CC allergenic of the 90 peptides tested.
 XX SQ Sequence 514 AA;
 Query Match 100.0%; Score 76; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LQFAKLTGFTLMGKG 15
 140 LQFAKLTGFTLMGKG 154
 Db
 RESULT 8
 AAR81586
 ID AAR81586 standard; Protein; 514 AA.
 XX AAR81586;
 AC AAR81586;
 DT 24-MAY-1996 (first entry)
 XX Cedar pollen allergen A.
 DE Cedar pollen allergen A.
 XX Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
 KW antibody; pollinosis; therapy; immunotherapy.
 XX Cryptomeria japonica.
 XX EP700929-A2.
 XX 13-MAR-1996.
 PD 08-SEP-1995; 95EP-0306295.
 XX 14-JUL-1995; 95UP-0200221.

PR 10-SEP-1994; 94UP-0242137.
 PR 14-JUL-1995; 95UP-0200204.
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX Hino K, Saito S, Taniguchi Y;
 DR WPI; 1996-140976/15.
 XX New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
 PT used for treating cedar pollinosis
 XX Claim 5; Page 29-30; 36pp; English.
 XX Synthetic peptides based on portions of cedar pollen allergens A
 CC (AAR81586) and B (AAR81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
 CC for T-cell recognition, and homologous peptides (AAR81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 XX SQ Sequence 514 AA;
 Query Match 100.0%; Score 76; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LQFAKLTGFTLMGKG 15
 140 LQFAKLTGFTLMGKG 154
 Db
 RESULT 9
 AAY25666
 ID AAY25666 standard; Protein; 514 AA.
 XX AAY25666;
 AC AAY25666;
 DT 30-SEP-1999 (first entry)
 XX Japanese cedar allergen 1076242 Cry j II precursor protein fragment.
 DE Major histocompatibility complex; class II; desensitizing; human;
 KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX Cedrus sp.
 OS Cedrus sp.
 XX MO9934826-A1.
 PD 15-JUL-1999.
 XX 11-JAN-1999; 99WO-GB00080.
 PR 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PA Kay AB, Larche M,
 XX Kay AB, Larche M,
 PI WPI; 1999-458255/38.
 DR Desensitizing patients to polypeptide allergens
 PT Example 6; Page 74; 117pp; English.
 XX

CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076242 Cry j II
 CC precursor.

CC Sequence 514 AA;

Query Match 100.0%; Score 76; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LQFAKLTGFTLMGKG 15
 DB 140 LQFAKLTGFTLMGKG 154

RESULT 10
 AAY25667

ID AAY25667 standard; protein; 514 AA.

AC AAY25667;

DT 30-SEP-1999 (first entry)

XX Japanese cedar allergen 1076241 Cry j II protein fragment.

XX Major histocompatibility complex; class II; desensitizing; human;
 KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KW chironomidae; spider; mite; housefly; fruit fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.

XX WO9934826-A1.

PD 15-JUL-1999.

XX 11-JAN-1999; 99WO-GB00080.

XX 21-SEP-1998; 98GB-0020474.

PR 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

PI Kay AB, Larche M;

DR WPI; 1999-45825/38.

XX Desensitizing patients to polypeptide allergens

XX Example 6; Page 74; 117pp; English.

CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for

CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076241 Cry j II.

CC Sequence 514 AA;

Query Match 100.0%; Score 76; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LQFAKLTGFTLMGKG 15
 DB 140 LQFAKLTGFTLMGKG 154

RESULT 11
 AAM27370

ID AAM27370 standard; peptide; 105 AA.

AC AAM27370;

DT 24-MAR-1998 (first entry)

XX Multi-epitope peptide used as immunotherapeutic agent #2.

XX Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

XX WO9732600-A1.

PD 12-SEP-1997.

XX 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

XX (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;

DR WPI; 1997-470495/43.

XX Peptide immuno-therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens

XX Claim 6; Page 31; 58pp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX Sequence 105 AA;

Query Match 88.2%; Score 67; DB 18; Length 105;
 Best Local Similarity 92.9%; Pred. No. 0.00013;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGK 14
 |||||
 DB 59 LQFAKLTGFTLMGR 72

RESULT 12
 AAW27371
 ID AAW27371 standard; peptide: 134 AA.

XX AAW27371;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #3.

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

XX Synthetic.

PN WO9732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Iwama A, Kuno K, Kume A, Sone T;

DR WPI; 1997-470495/43.

PT Peptide immunotherapeutic agent to treat allergic diseases -
 contains multi-epitope peptide containing T cell epitope regions

PT from different allergens

PS Claim 6; Page 32; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX Sequence 134 AA;

Query Match 88.2%; Score 67; DB 18; Length 134;

Best Local Similarity 92.9%; Pred. No. 0.00017;

Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGK 14
 |||||
 DB 59 LQFAKLTGFTLMGR 72

RESULT 13
 AAW12535
 ID AAW12535 standard; peptide: 14 AA.

XX AAW12535;

DT 30-APR-1997 (first entry)

DE Japanese cedar pollen allergen-derived peptide 4.

KW Japanese cedar; sugi; pollen; pollinosis; pollenosis; allergen;
 KW allergy; T-cell epitope; TCR; T-cell receptor; activation;
 KW immune tolerance; CryJ).

OS Crytomeria japonica.

PN DP08333391-A.

PD 17-DEC-1996.

PF 18-JUL-1995; 95JP-0181438.

PR 07-APR-1995; 95JP-0082519.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

PI (SANY) SANKYO CO LTD.

DR WPI; 1997-095487/09.

PT Peptide allergen derived from Japanese cedar pollen - causes T cell
 PT response specific to cedar pollen, for treatment of pollinosis

PS Claim 5; Page 19; 21pp; Japanese.

XX The present sequence is one of 24 claimed peptides which were
 CC synthesised based on Japanese cedar pollen sequences. This peptide
 CC was shown to have Japanese cedar pollen antigen T cell epitope
 CC activity by using T cells isolated from a cedar pollinosis patient.
 CC The peptide produces little or no anaphylaxis.

XX Sequence 14 AA;

Query Match 85.5%; Score 65; DB 18; Length 14;

Best Local Similarity 100.0%; Pred. No. 3.2e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMGK 13
 |||||
 DB 2 LQFAKLTGFTLMGK 14

RESULT 14
 AAR81581
 ID AAR81581 standard; peptide: 17 AA.

XX AAR81581;

DT 24-MAY-1996 (first entry)

DE Cedar pollen allergen peptide 9 (T-cell epitope).

KW Cedar; Crytomeria japonica; pollen; allergen; immunoglobulin E;

KW IgE; T-cell epitope; antibody; pollinosis; therapy; immunotherapy.

XX Synthetic.

PN EP700929-A2.

PD 13-MAR-1996.

PF 08-SEP-1995; 95EP-0306295.

PR 14-JUL-1995; 95JP-0200221.

PR 10-SEP-1994; 94JP-0242137.

PR 14-JUL-1995; 95JP-0200204.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

PI Hino K, Saito S, Taniguchi Y;

DR WPI; 1996-140976/15.

PT New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IGE antibodies,
 PT used for treating cedar pollinosis

XX Claim 4; Page 28; 36pp; English.

XX Synthetic peptides based on portions of cedar pollen allergens A
 CC (AAR81586) and B (AAR81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IGE
 CC antibodies. 6 Peptides (AAR81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
 CC for T-cell recognition, and homologous peptides (AAR81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.

XX Sequence 17 AA;

Query Match 85.5%; Score 65; DB 17; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3.9e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQFAKLTGFTLMG 13
 Db 5 LQFAKLTGFTLMG 17

RESULT 15

AAB69104
 ID AAB69104 standard; Protein; 96 AA.

AC AAB69104;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:13.

XX Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KM T cell epitope; anti-sugipollinosis.

XX Cryptomeria japonica.

OS Synthetic.

PN JP2000327699-A.

PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.

PR 15-MAR-1999; 99JP-0068316.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.

DR WPI; 2001-185061/19.

XX Novel peptide and its use -

PT Claim 8; Page 40; 75pp; Japanese.

XX The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (I) has a formula of:

CC alpha1-alpha3-alpha4-alpha5-alpha6-alpha7
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.

CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.

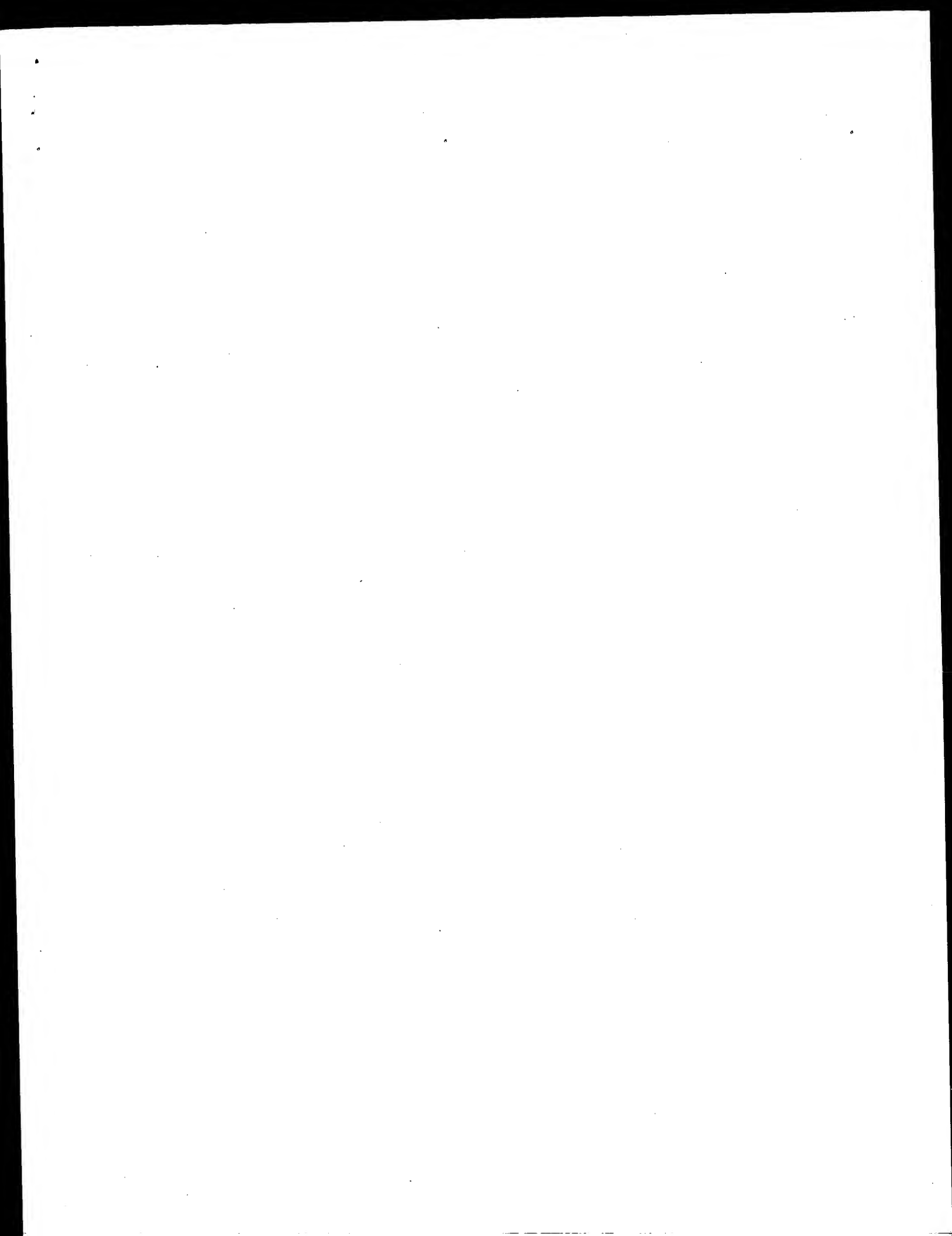
XX Sequence 96 AA;

Query Match 82.9%; Score 63; DB 22; Length 96;
 Best Local Similarity 92.3%; Pred. No. 0.00065;

Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QFAKLTGFTLMGX 14
 Db 84 QFAKLTGFTLMGR 96

Search completed: April 20, 2003, 13:06:10
 Job time: 19.1974 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds
(without alignments) 109.838 Million cell updates/sec

Title: US-09-142-524d-120

Sequence: 1 GIDIPASKNFHLQKN 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	15	17	AA897907
2	79	100.0	15	19	AA857763
3	79	100.0	460	16	AA869791
4	79	100.0	514	15	AA853690
5	79	100.0	514	16	AA874333
6	79	100.0	514	16	AA869792
7	79	100.0	514	17	AA835999
8	79	100.0	514	17	AA81586
9	79	100.0	514	20	AA85666
10	79	100.0	514	20	AA85667

11	73	92.4	17	19	AA80347	Sugi allergen prot
12	73	92.4	19	22	AA884115	Amio acid sequenc
13	73	92.4	42	22	AA884105	Immunomodulatory p
14	73	92.4	80	18	AAW27369	Multi-epitope pep
15	73	92.4	105	18	AAW27370	Multi-epitope pep
16	73	92.4	134	18	AAW27371	Multi-epitope pep
17	72	91.1	74	21	AA823905	Artificial sequenc
18	72	91.1	94	21	AA823906	Artificial sequenc
19	72	91.1	95	21	AA823897	Artificial sequenc
20	72	91.1	95	22	AA823901	Artificial sequenc
21	72	91.1	95	22	AA869098	Artificial sequenc
22	72	91.1	96	22	AA869104	Artificial sequenc
23	72	91.1	96	22	AA869105	Artificial sequenc
24	72	91.1	97	22	AA869105	Artificial sequenc
25	72	91.1	99	21	AA823902	Cedar pollen aller
26	72	91.1	99	21	AA823907	Cedar pollen aller
27	72	91.1	99	22	AA869099	Artificial sequenc
28	72	91.1	214	22	AA869120	Artificial sequenc
29	69	87.3	13	22	AA869101	Artificial sequenc
30	69	87.3	17	17	AA81582	Cryptomeria japoni
31	63	79.7	12	18	AAW12548	Cryptomeria japoni
32	62	78.5	20	19	AAW42181	Cryptomeria japoni
33	62	78.5	514	17	AAW04346	Cedar pollen aller
34	62	78.5	514	19	AAW42122	Cedar pollen aller
35	61	77.2	33	19	AAW80340	T-cell epitope pep
36	61	77.2	33	19	AAW80341	Chamaecyparis obtu
37	59	74.7	11	18	AAW12550	Japanese cypress p
38	59	74.7	12	18	AAW12549	Sugi allergen prot
39	59	74.7	12	22	AAW69108	Sugi allergen prot
40	59	74.7	33	19	AAW80339	Sugi allergen prot
41	59	74.7	35	19	AAW80342	Sugi allergen prot
42	59	74.7	35	19	AAW80343	Sugi allergen prot
43	59	74.7	35	19	AAW80344	Sugi allergen prot
44	59	74.7	47	19	AAW80350	Sugi allergen prot
45	59	74.7	47	19	AAW80351	Sugi allergen prot

ALIGNMENTS

RESULT 1
AA897907 standard; peptide; 15 AA.
AA897907;
16-AUG-1996 (first entry)
Japan cedar pollen mature allergen Cry j II amino acids 181-195.
Allergen: epitope; overlapping peptide; Cry j II; cedar pollen;
Sugi pollinosis; diagnosis; treatment.
Cryptomeria japonica.
JP08047392-A.
20-FEB-1996.
07-NOV-1994; 94UP-0297840.
26-MAY-1994; 94UP-0134868.
05-NOV-1993; 93UP-0276773.
(MEIP) MEIJI MILK PROD CO LTD.
WPI; 1996-166249/17.
Japan cedar pollen allergen Cry j II epitope - comprises at least part of specified 460 amino acid protein
Claim 8; Fig 4; 17pp; Japanese.

CC AAR97871-R97960 are overlapping peptides used for the epitope mapping
 CC of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic
 CC peptides of it are useful in the diagnosis, prevention and treatment
 CC of Sugi pollinosis, the allergic reaction to Japan cedar pollen.
 CC Significant regions of the allergen were identified using the
 CC overlapping peptides of the full epitope derived from a Cry j II
 CC antigen-specific T cell line. Amino acids 66-80 (AAR97884) and 186-200
 CC (R978908) of the full mature 460 amino acid allergen are the most
 CC allergenic of the 90 peptides tested.

CC Sequence 15 AA;

Query Match 100.0%; Score 79; DB 17; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.1e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHLQKN 15
 DB 1 GIDIFASKNFHLQKN 15

RESULT 2
 AAM57763
 ID AAM57763 standard; peptide; 15 AA.

AC AAM57763;

DT 17-SEP-1998 (first entry)

DE Residues 181-195 of Cry j 2.

KW Cry j 2; Japanese cedar pollen antigen; allergy; immunotherapy;

KM HLA class II molecule.

OS Cryptomeria japonica.

PN W09820902-A1.

PD 22-MAY-1998.

PF 12-NOV-1997; 97MO-JP04129.

PR 13-NOV-1996; 96JP-0302053.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Kino K, Kume A, Sone T;

PI WPI; 1998-297617/26.

PT Peptides derived from Japanese cedar pollen antigens are
 PT immunotherapeutic agents - useful for allergy treatment and typing

PS HLA class II molecules in allergy sufferers

PS Claim 12; Page 31; 50pp; Japanese.

CC This sequence represents residues 181-195 of the Cry j 2 protein, and
 CC is a peptide of the invention. The peptides are derived from Japanese
 CC cedar pollen antigens, and are used as immunotherapeutic agents in the
 CC treatment of allergy. The peptides can be used for identification and
 CC typing of the particular HLA class II molecules in an allergy sufferer,
 CC and also for peptide immunotherapy of an allergy. Using these peptides
 CC the immunotherapy can be targeted more specifically to the requirements
 CC of the individual patient, allowing more effective treatment of an
 CC allergy, including those patients for whom treatment with a conventional
 CC immunotherapeutic agent is ineffective.

XX Sequence 15 AA;

Query Match 100.0%; Score 79; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.1e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHLQKN 15
 DB 1 GIDIFASKNFHLQKN 15

RESULT 3
 AAR69791
 ID AAR69791 standard; Protein; 460 AA.

AC AAR69791;

DT 27-SEP-1995 (first entry)

DE Japonicum allergen residues 55-514.

KW Japonicum allergen; residues 55-514; induced histamine release;

KM antiallergic peptide; IgE cross-linking inhibition.

OS Japonicum sp.

PN W09502412-A.

PD 26-JAN-1995.

PF 15-JUL-1994; 94MO-JP01164.

PR 16-JUL-1993; 93JP-0177008.

PR 01-SEP-1993; 93JP-0217725.

PR 07-APR-1994; 94JP-0069336.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Kino K, Kohno Y, Komiya N, Sone T;

PI WPI; 1995-067159/09.

DR N-PSDB; AAO84044.

PT Peptide antiallergic agent - inhibits cross-linking of allergen
 PT with IgE antibody

PS Disclosure; Pages 26-27; 46pp; Japanese.

CC AAO84044 encodes AAR69791 Japonicum allergen residues 55-514, from
 CC which the antiallergic peptides AAR69845-R69809 were derived.
 CC The peptides ability to inhibit the cross-linking of an allergen,
 CC to an IgE antibody can be used in the prevention and treatment of
 CC allergic diseases.

XX Sequence 460 AA;

Query Match 100.0%; Score 79; DB 16; Length 460;
 Best Local Similarity 100.0%; Pred. No. 2.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIDIFASKNFHLQKN 15
 DB 181 GIDIFASKNFHLQKN 195

RESULT 4
 AAR53690
 ID AAR53690 standard; Protein; 514 AA.

AC AAR53690;

DT 01-FEB-1995 (first entry)

DE Japanese cedar pollen allergen Cry j II.

KW Cedar pollinosis; diagnostic.

OS Cryptomeria japonica.

XX

PN W09411512-A.
 XX
 PD 26-MAY-1994.
 XX
 PF 12-NOV-1993; 93WO-US11000.
 XX
 PR 12-NOV-1992; 92US-0975179.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Brauer A, Kuo M, Pollock J, Yeung S;
 XX
 DR WPI; 1994-183513/22.
 DR N-PSDB; AAQ65048.
 XX
 PT Allergenic Cry j II protein and fragments from Japanese cedar
 PT pollen - used to diagnose, treat and prevent Japanese cedar
 XX pollinosis
 PS Claim 2; Fig 4; 89pp; English.
 CC
 CC The sequence is of a Japanese cedar pollen allergen Cry j
 CC II. The protein and its fragments can be used for diagnosis and
 CC treatment of Japanese cedar pollinosis and to identify similar
 CC sequences in other plants.
 CC See also AAR33692-6.
 CC
 SQ Sequence 514 AA;
 XX
 XX
 Query Match 100.0%; Score 79; DB 15; Length 514;
 Best Local Similarity 100.0%; Pred. No. 3.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GIDIFASKNFHLOKN 15
 |||||
 Db 235 GIDIFASKNFHLOKN 249
 |||||
 RESULT 5
 AAR74333
 ID AAR74333 standard; Protein; 514 AA.
 XX
 AC AAR74333;
 XX
 XX 01-NOV-1995 (first entry)
 DT
 XX Japanese cedar pollen allergen.
 DE
 XX Japanese cedar pollen allergen.
 KW Japanese cedar; pollen; allergen; allergy; therapy; diagnostic;
 KM desensitizer.
 XX
 OS Cryptomeria japonica.
 OS
 PN EP655500-A.
 XX
 PD 31-MAY-1995.
 PD
 PF 03-NOV-1994; 94EP-0308117.
 PF
 PR 27-DEC-1993; 93JP-0346814.
 PR 05-NOV-1993; 93JP-0299151.
 PR 20-DEC-1993; 93JP-0344596.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 PI Kurimoto M, Namba M, Torigoe K;
 XX
 DR WPI; 1995-195588/26.
 DR N-PSDB; AAQ90156.
 XX
 PT New Japanese cedar pollen allergen polypeptide - and DNA coding for it,
 PT useful for treatment and diagnosis of cedar pollen allergy
 XX

PS Claim 5; Page 26-28; 41pp; English.
 XX
 CC The gene encoding an allergen of Japanese cedar pollen was isolated
 CC by PCR amplification using primers based on portions of the allergen
 CC protein. The gene was used for recombinant allergen production in
 CC E. coli (vector plasmid PKK-223-3).
 CC
 SQ Sequence 514 AA;
 XX
 XX
 Query Match 100.0%; Score 79; DB 16; Length 514;
 Best Local Similarity 100.0%; Pred. No. 3.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GIDIFASKNFHLOKN 15
 |||||
 Db 235 GIDIFASKNFHLOKN 249
 |||||
 RESULT 6
 AAR69792
 ID AAR69792 standard; Protein; 514 AA.
 XX
 AC AAR69792;
 XX
 DT 27-SEP-1995 (first entry)
 DT
 XX Japonicum allergen.
 DE Japonicum allergen.
 XX
 KW Japonicum allergen; induced histamine release; antiallergic peptide;
 KW Ige cross-linking inhibition.
 XX
 OS Japonicum sp.
 OS
 PN W09502412-A.
 XX
 PD 26-JAN-1995.
 PD
 PF 15-JUL-1994; 94WO-JP01164.
 PF
 PR 16-JUL-1993; 93JP-0177008.
 PR 01-SEP-1993; 93JP-0217725.
 PR 07-APR-1994; 94JP-0069336.
 XX
 PA (MEIP) MEIJI MILK PROD CO LTD.
 XX
 PI Kino K, Kohno Y, Komiyama N, Sone T;
 XX
 DR WPI; 1995-067159/09.
 DR N-PSDB; AAQ84045; AAQ84046.
 XX
 PT Peptide antiallergic agent - inhibits cross-linking of allergen
 PT with Ige antibody
 XX
 PS Example 3; Pages 27-28; 46pp; Japanese.
 XX
 CC AAQ84045 encodes AAR69792 Japonicum allergen, from which the
 CC antiallergic peptides AAR69845-R69809 were derived. The peptides ability
 CC to inhibit the cross-linking of an allergen, to an Ige antibody can be
 CC used in the prevention and treatment of allergic diseases.
 CC
 SQ Sequence 514 AA;
 XX
 XX
 Query Match 100.0%; Score 79; DB 16; Length 514;
 Best Local Similarity 100.0%; Pred. No. 3.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GIDIFASKNFHLOKN 15
 |||||
 Db 235 GIDIFASKNFHLOKN 249
 |||||
 RESULT 7
 AAR93599

ID AAR93599 standard; Protein: 514 AA.
 AC AAR93599;
 DT 16-AUG-1996 (first entry)
 DE Japan cedar pollen Cry j II allergen.
 DE Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
 KM Sugi pollinosis; diagnosis; treatment.
 OS Cryptomeria japonica.
 FH Key Location/Qualifiers
 FT Peptide 1..54
 FT /*label= sig_peptide
 FT 55..514
 FT /*label= mat_protein
 PN JP08047392-A.
 PD 20-FEB-1996.
 PF 07-NOV-1994; 94JP-0297840.
 PR 26-MAY-1994; 94JP-0134868.
 PR 05-NOV-1993; 93JP-0276773.
 PA (MEIRI) MEIRI MILK PROD CO LTD.
 DR MPI; 1996-166249/17.
 DR N-PSDB; AAT18102.
 XX Japan cedar pollen allergen Cry j II epitope - comprises at least
 PT part of specified 460 amino acid protein
 PS Claim 1; Page 10-11; 17pp; Japanese.
 CC AAR93599 is a Japan cedar pollen Cry j II allergen which is useful
 CC in the diagnosis, prevention and treatment of Sugi pollinosis;
 CC the allergic reaction to Japan cedar pollen. Significant regions of
 CC the allergen were identified using overlapping peptides of the full
 CC epitope derived from a Cry j II antigen-specific T cell line
 CC (see AAR97871-R97960). Amino acids 66-80 (AAR97884) and 186-200
 CC (R978908) of the full mature 460 amino acid allergen are the most
 CC allergenic of the 90 peptides tested.
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 79; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 3.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GIDIFASKNFHLQKN 15
 DB 235 GIDIFASKNFHLQKN 249
 RESULT 8
 AAR81586
 ID AAR81586 standard; Protein: 514 AA.
 AC AAR81586;
 DT 24-MAY-1996 (first entry)
 DE Cedar pollen allergen A.
 DE Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
 KM antibody; pollinosis; therapy; immunotherapy.
 OS Cryptomeria japonica.

PN EP700929-AZ.
 PD 13-MAR-1996.
 PF 08-SEP-1995; 95SEP-0306295.
 PR 14-JUL-1995; 95JP-0200221.
 PR 10-SEP-1994; 94JP-0242137.
 PR 14-JUL-1995; 95JP-0200204.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 DR Hino K, Saito S, Taniguchi Y;
 DR MPI; 1996-140976/15.
 PT New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
 PT used for treating cedar pollinosis
 PS Claim 5; Page 29-30; 36pp; English.
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (AAR81586) and B (AAR81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 Peptides (AAR81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
 CC for T-cell recognition, and homologous peptides (AAR81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 SQ Sequence 514 AA;
 Query Match 100.0%; Score 79; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 3.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GIDIFASKNFHLQKN 15
 DB 235 GIDIFASKNFHLQKN 249
 RESULT 9
 AAY25666
 ID AAY25666 standard; Protein: 514 AA.
 AC AAY25666;
 DT 30-SEP-1999 (first entry)
 DE Japanese cedar allergen 1076242 Cry j II precursor protein fragment.
 DE Major histocompatibility complex; class II; desensitizing; human;
 KM allergen; grass; tree; weed; pollen; fungi; mold; food; insect; sting;
 KM chironomidae; spider; mite; housefly; fruit fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larva; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 OS Cedrus sp.
 PN WO9934826-A1.
 PD 15-JUL-1999.
 PF 11-JAN-1999; 99WO-GB00080.
 PR 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PI Kay AB, Larche M;

XX WPI, 1999-458255/38.
XX

XX Desensitizing patients to polypeptide allergens
XX

XX Example 6; Page 74; 117pp; English.
XX

CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC Tenbrio molitor beetle, bee moth larvae, mealworm, cockroach, larvae of
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a Japanese cedar (Cedrus sp.) allergen 1076242 Cry j II
CC precursor.
XX

XX Sequence 514 AA;
XX

Query Match 100.0%; Score 79; DB 20; Length 514;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIDIFASKNFHLQKN 15
DB 235 GIDIFASKNFHLQKN 249

RESULT 10

AA25667
ID AAY25667 standard; protein; 514 AA.

XX AAY25667;
XX

XX 30-SEP-1999 (first entry)
XX

XX Japanese cedar allergen 1076241 Cry j II protein fragment.
XX

XX Major histocompatibility complex; class II; desensitizing; human;
XX allergen; grass; tree; weed; pollen; fungi; mould; food; insect; stinging;
XX chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX cockroach; beetle; dog; horse; pig; sheep; rabbit; rat; guinea pig;
XX mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX

XX Cedrus sp.
XX

XX WO9934826-A1.
XX

XX 15-JUL-1999.
XX

XX 11-JAN-1999; 99WC-GB00080.
XX

XX 21-SEP-1998; 98GB-0020474.
XX

XX 09-JAN-1998; 98GB-0000445.
XX

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX

XX Kay AB, Larche M;
XX

XX WPI, 1999-458255/38.
XX

XX Desensitizing patients to polypeptide allergens
XX Example 6; Page 74; 117pp; English.
XX

XX This invention describes a novel method of desensitizing a patient to a
XX polypeptide allergen and comprises administering to the patient a peptide
XX derived from the allergen where restriction to a MHC Class II molecule
XX possessed by the patient can be demonstrated for the peptide and the
XX peptide is able to induce a late phase response in an individual who
XX possesses the MHC Class II molecule. The methods can be used for
XX desensitizing patients to allergens present in e.g. grass, tree and weed
XX (including ragweed) pollens, fungi and moulds, foods, stinging insects,
XX the chironomidae (non-biting midges), spiders and mites, housefly, fruit
XX fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
XX Tenbrio molitor beetle, bee moth larvae, mealworm, cockroach, larvae of
XX sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
XX produce immunological vaccines which may be used to prevent and/or treat
XX conditions involving hypersensitivity to allergens. This sequence
XX represents a Japanese cedar (Cedrus sp.) allergen 1076241 Cry j II.
XX

XX Sequence 514 AA;
XX

Query Match 100.0%; Score 79; DB 20; Length 514;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIDIFASKNFHLQKN 15
DB 235 GIDIFASKNFHLQKN 249

RESULT 11

AAW80347
ID AAW80347 standard; peptide; 17 AA.

XX AAW80347;
XX

XX 11-JAN-1999 (first entry)
XX

XX Sugi allergen protein Cryj2 derived epitope for T cells.
XX

XX T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;
XX sugi-pollinosis; allergic reaction; pollen.
XX

XX Synthetic.
XX

XX JP10259198-A.
XX

XX 29-SEP-1998.
XX

XX 22-DEC-1997; 97JP-0353448.
XX

XX 24-DEC-1996; 96JP-0343441.
XX

XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX (SANY) SANKYO CO LTD.
XX

XX WPI, 1998-577037/49.
XX

XX A linked T cell epitope peptide - used for the treatment of
XX sugi-pollinosis
XX

XX Claim 7; Page 18; 21pp; Japanese.
XX

XX AAW80339-58 represent epitopes for T cells, derived from the sugi
XX allergen proteins Cryj1 (AAW80339-44, AAW80350-53 and AAW80356-58) and
XX Cryj2 (AAW80345-49 and AAW80354-55). The peptides are useful for the
XX treatment of sugi-pollinosis, an allergic reaction of the body to
XX pollen.
XX

XX Sequence 17 AA;
XX

Query Match 92.4%; Score 73; DB 19; Length 17;
Best Local Similarity 100.0%; Pred. No. 9.2e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IDIFASKNFHLQKN 15
 DB 1 IDIFASKNFHLQKN 14

RESULT 12

AAB84115
 ID AAB84115 standard; peptide; 19 AA.

AC AAB84115;
 XX

DT 06-AUG-2001 (first entry)

DE Amino acid sequence of a pollen antigen.

KW Immunomodulatory peptide; immune response; autoimmune disease; allergy;
 asthma; host-versus-graft rejection; T cell; anergy; apoptosis.

OS Cryptomeria japonica.

PN WO200136448-A2.

PD 25-MAY-2001.

PF 27-OCT-2000; 2000WO-US41646.

PR 27-OCT-1999; 99US-0161734.

PA (CELS-) CEL-SCI CORP.

PI Zimmerman DH;

DR WPI; 2001-374498/39.

XX Novel immunomodulatory peptide construct useful for modulating an
 PT inappropriate immune response in an individual at risk for autoimmune
 PT disease, allergic reactions, asthma or host-graft or graft-host disease

PS Claim 2; Page 27; 55pp; English.

XX The specification describes an immunomodulatory peptide construct.
 CC The immunomodulatory peptide comprises a first peptide associated with
 CC autoimmune disease, allergy or asthma, or host-versus-graft rejection
 CC and which will bind to an antigen receptor on a set or subset of T cells,
 CC linked to a second immune modulating peptide which will cause a directed
 CC immune response by the set or subset of T cells to which the first
 CC peptide is attached. Alternatively, the second peptide will bind to a
 CC T cell receptor site on the surface of the T cell which will cause the
 CC set or subset of T cells to initiate, but not complete, an immune
 CC response to cause the set or subset of T cells to undergo anergy and
 CC apoptosis. The immunomodulatory peptides are useful for eliminating a set
 CC or subset of T cells involved in autoimmune response. They are useful for
 CC the treatment of autoimmune disease, allergic reactions, asthma or
 CC host-graft or graft-host rejection. The immunomodulatory peptides are
 CC also useful for interrupting an autoimmune disease associated pathway
 CC necessary to complete T cell activation. The present peptide is used to
 CC construct immunomodulatory peptides of the invention, and is
 CC representative of the first peptide.

SO Sequence 19 AA;

Query Match 92.4%; Score 73; DB 22; Length 19;

Best Local Similarity 100.0%; Pred. No. 1e-06; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0;

OY 2 IDIFASKNFHLQKN 15
 DB 1 IDIFASKNFHLQKN 14

RESULT 13

AAB84105
 ID AAB84105 standard; peptide; 42 AA.

AC AAB84105;
 XX

DT 06-AUG-2001 (first entry)

DE Immunomodulatory peptide used to treat allergic responses.

KW Immunomodulatory peptide; immune response; autoimmune disease; allergy;
 asthma; host-versus-graft rejection; T cell; anergy; apoptosis.

OS Synthetic.
 Cryptomeria japonica.

FT Key Location/Qualifiers

FT Peptide 1..17 /note= "LFA-3 peptide"

FT Peptide 18..23 /note= "linker"

FT Peptide 24..42 /note= "pollen antigen"

PN WO200136448-A2.

PD 25-MAY-2001.

PF 27-OCT-2000; 2000WO-US41646.

PR 27-OCT-1999; 99US-0161734.

PA (CELS-) CEL-SCI CORP.

PI Zimmerman DH;

DR WPI; 2001-374498/39.

XX Novel immunomodulatory peptide construct useful for modulating an
 PT inappropriate immune response in an individual at risk for autoimmune
 PT disease, allergic reactions, asthma or host-graft or graft-host disease

PS Claim 4; Page 27; 55pp; English.

XX The specification describes an immunomodulatory peptide construct.
 CC The immunomodulatory peptide comprises a first peptide associated with
 CC autoimmune disease, allergy or asthma, or host-versus-graft rejection
 CC and which will bind to an antigen receptor on a set or subset of T cells,
 CC linked to a second immune modulating peptide which will cause a directed
 CC immune response by the set or subset of T cells to which the first
 CC peptide is attached. Alternatively, the second peptide will bind to a
 CC T cell receptor site on the surface of the T cell which will cause the
 CC set or subset of T cells to initiate, but not complete, an immune
 CC response to cause the set or subset of T cells to undergo anergy and
 CC apoptosis. The immunomodulatory peptides are useful for eliminating a set
 CC or subset of T cells involved in autoimmune response. They are useful for
 CC the treatment of autoimmune disease, allergic reactions, asthma or
 CC host-graft or graft-host rejection. The immunomodulatory peptides are
 CC also useful for interrupting an autoimmune disease associated pathway
 CC necessary to complete T cell activation. The present sequence represents
 CC an immunomodulatory peptide used to treat allergic responses.

SO Sequence 42 AA;

Query Match 92.4%; Score 73; DB 22; Length 42;

Best Local Similarity 100.0%; Pred. No. 2.6e-06; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0;

OY 2 IDIFASKNFHLQKN 15
 DB 24 IDIFASKNFHLQKN 37

RESULT 14

AAW27369 standard; peptide; 80 AA.

AC AAW27369;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #1.

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN WO9732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;

DR WPI; 1997-470495/43.

PT Peptide immuno:therapeutic agent to treat allergic diseases -
contains multi-epitope peptide containing T cell epitope regions
from different allergens

PS Claim 6; Page 31; 58pp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as
a new immunotherapeutic agent. It comprises T cell epitope regions from 2
or more different allergens (preferably linked via arginine or lysine
dimers), where the T cell epitope regions have a positivity index
greater than 100 as measured in a patient group responding to the
allergen; have at least 70% reactivity with lymphocytes from patients
responding to the allergen; and are not reactive with immunoglobulin E
(IgE) antibodies from patients responsive to the allergen. The agent can
be used to prevent and treat a wide variety of allergic diseases, e.g. by
desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

SQ Sequence 80 AA;

Query Match 92.4%; Score 73; DB 18; Length 80;

Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IDIFASKNPHLOKN 15

DB 31 IDIFASKNPHLOKN 44

RESULT 15

AAW27370 standard; peptide; 105 AA.

AC AAW27370;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #2.

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

EN WO9732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;

DR WPI; 1997-470495/43.

PT Peptide immuno:therapeutic agent to treat allergic diseases -
contains multi-epitope peptide containing T cell epitope regions
from different allergens

PS Claim 6; Page 31; 58pp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as
a new immunotherapeutic agent. It comprises T cell epitope regions from 2
or more different allergens (preferably linked via arginine or lysine
dimers), where the T cell epitope regions have a positivity index
greater than 100 as measured in a patient group responding to the
allergen; have at least 70% reactivity with lymphocytes from patients
responding to the allergen; and are not reactive with immunoglobulin E
(IgE) antibodies from patients responsive to the allergen. The agent can
be used to prevent and treat a wide variety of allergic diseases, e.g. by
desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

SQ Sequence 105 AA;

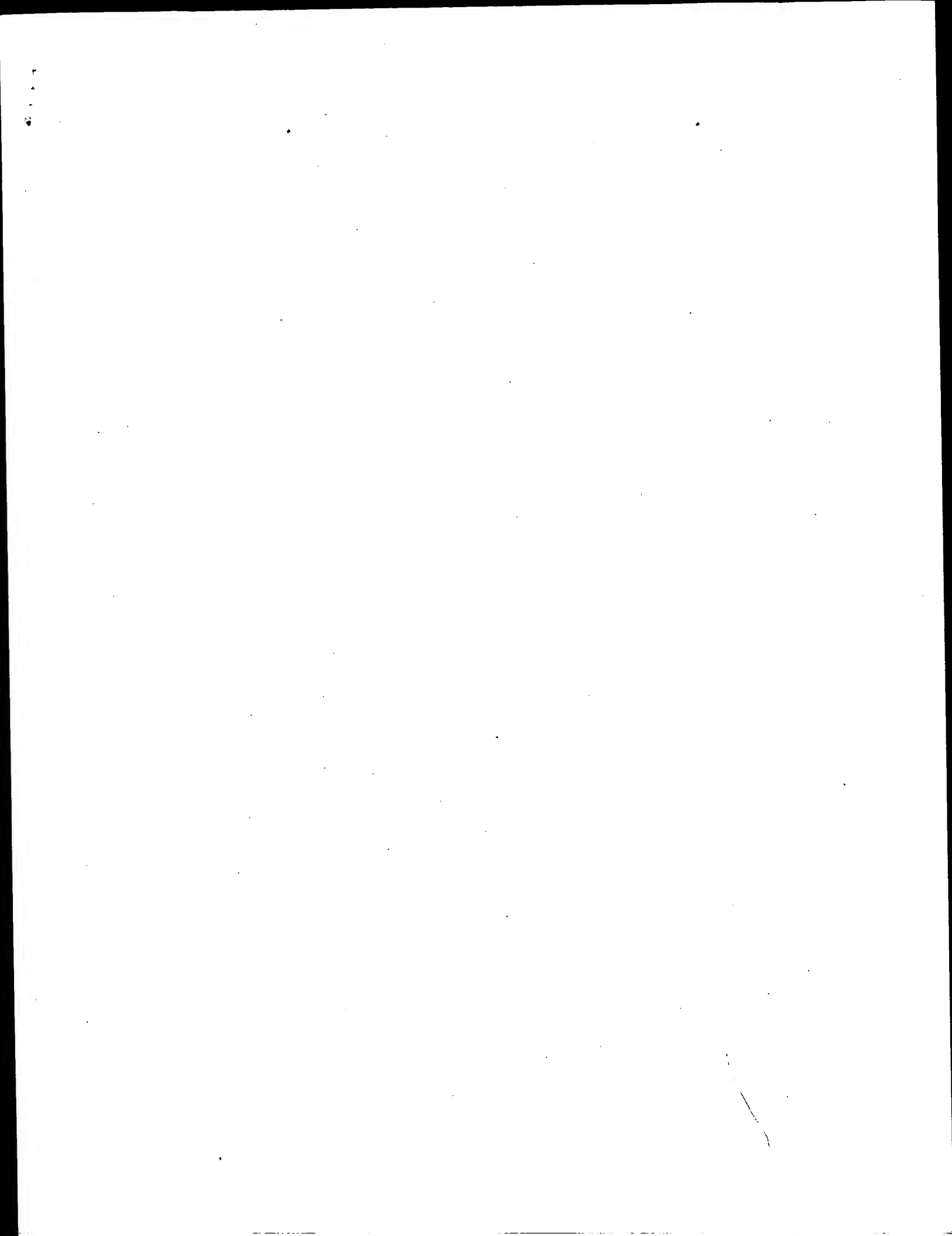
Query Match 92.4%; Score 73; DB 18; Length 105;

Best Local Similarity 100.0%; Pred. No. 7.3e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IDIFASKNPHLOKN 15

DB 31 IDIFASKNPHLOKN 44

Search completed: April 20, 2003, 13:06:11
Job time : 19.1974 secs



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 2.92105 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524d-152

Perfect score: 67

Sequence: 1 LSDISLKLTKSGKIAS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	514	MPA2_CRYJA	P43212 cryptomeria
2	43	64.2	251	ZNUC_ECOLI	P52648 escherichia
3	42	62.7	255	PECE_ECOLI	P15031 escherichia
4	38	56.7	145	VP12_AGRPU	O44433 agrobacteri
5	38	56.7	316	Y248_CAEEL	O76387 caenorhabdi
6	38	56.7	530	U121_HSVB	P28972 equine herp
7	37	55.2	231	YTRF_BACSU	O34392 bacillus su
8	37	55.2	232	YHCG_BACSU	P54591 bacillus su
9	37	55.2	311	Y467_MYCB	P47705 mycoplasma
10	37	55.2	454	PR11_SCHPO	O14215 schizosacch
11	37	55.2	504	PD43_MOUSE	P27773 mus musculu
12	36	53.7	229	CG18_YEAST	P38794 saccharomyc
13	36	53.7	250	MNTB_BACSU	O34338 bacillus su
14	36	53.7	365	CYSA_ECOLI	P16676 escherichia
15	36	53.7	365	CYSA_ECOLI	P40860 salmonella
16	36	53.7	365	CYSA_ECOLI	P40860 salmonella
17	36	53.7	507	MTPS_PROST	O57538 haemophilus
18	35	52.2	552	Y664_HAERH	P08874 bacillus su
19	35	52.2	96	ABRB_BACSU	P36925 ovine aries
20	35	52.2	101	ILB_SHEEP	P57403 buchnera ap
21	35	52.2	238	ZNUC_BUCAI	P45518 streptomyce
22	35	52.2	264	FTSQ_STRCO	P95788 streptococc
23	35	52.2	292	ATPG_STRMU	P75110 mycoplasma
24	35	52.2	339	Y467_MYCPN	P71827 mycobacteri
25	35	52.2	422	PUR2_MYCPN	P32827 mycobacteri
26	35	52.2	466	ECOS1_YEAST	P40086 saccharomyc
27	35	52.2	486	COXM_YEAST	P75059 mycoplasma
28	35	52.2	560	POTA_MYCPN	O99968 mus musculu
29	35	52.2	918	ABG5_MOUSE	P27801 saccharomyc
30	35	52.2	944	PEP3_YEAST	O9446 rhizobium l
31	35	52.2	948	GLND_RHITLV	O33244 rhizobium t
32	35	52.2	1186	GLND_RHITLV	O45710 bacillus th
33	35	52.2	1391	LYS2_CANAL	O12572 candida alb

34	34.5	51.5	722	1	MESD_LEBME	Q10418 leuconostoc
35	34	50.7	192	1	YF21_ARCPV	O28751 archaeeoglob
36	34	50.7	196	1	UHPA_ECOLI	P10940 escherichia
37	34	50.7	247	1	VGS_BPPI	P22588 bacteriopho
38	34	50.7	255	1	TAUB_ECOLI	O47538 escherichia
39	34	50.7	266	1	HMUV_YERPE	O56993 yeastsina pe
40	34	50.7	355	1	CYSA_SYNY3	P74548 synechocyst
41	34	50.7	363	1	YAS8_CHUPN	O94641 chlamydia p
42	34	50.7	365	1	PCP_SYMP	P51874 symbiodiniu
43	34	50.7	423	1	ENO_METVA	O60173 methanococc
44	34	50.7	435	1	ASPG_MERTH	O26802 methanobact
45	34	50.7	493	1	RBSA_HAERH	P44735 haemophilus

ALIGNMENTS

```

RESULT 1
MPA2_CRYJA          STANDARD;          PRT;          514 AA.
ID      P43212;
AC      P43212;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)
DE      (Major pollen allergen Cry j 2) (Cry j II).
OS      Cryptomeria japonica (Japanese cedar).
OC      Bkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OX      NCBI_TaxID=3369;
RN      [1]
RP      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC      TISSUE=Pollen;
RX      MEDLINE=95010777; PubMed=7926035;
RA      Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,
RA      Usui M., Kurimoto M.;
RT      "Molecular cloning of the second major allergen, Cry j II, from
RT      Japanese cedar pollen."
RL      FEBS Lett. 353:124-128 (1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Pollen;
RX      MEDLINE=94271186; PubMed=8002972;
RA      Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.;
RT      "CDNA cloning and expression of Cry j II the second major allergen of
RT      Japanese cedar pollen."
RL      Biochem. Biophys. Res. Commun. 201:1021-1028(1994).
RN      [3]
RP      SEQUENCE OF 55-64.
RX      MEDLINE=90342988; PubMed=2382797;
RA      Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matuhashi T.;
RT      "Identification of the second major allergen of Japanese cedar
RT      pollen."
RL      Allergy 45:309-312(1990).
RN      [4]
RP      CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
RP      galactosiduronic linkages in pectate and other galacturonans.
RN      [5]
RP      SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).
RN      [6]
RP      SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
RP      (POLYGALACTURONASES).
RN      [7]
RP      This SWISS-PROT entry is copyright. It is produced through a collaboration
RP      between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP      the European Bioinformatics Institute. There are no restrictions on its
RP      use by non-profit institutions as long as its content is in no way
RP      modified and this statement is not removed. Usage by and for commercial
RP      entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP      or send an email to license@isb-sib.ch).
RN      [8]
RP      EMBL; D37765; BAA07021.1; -
RN      EMBL; D29772; BAA06172.1; -
RN      HSSP; P26509; 1BHE.
RN      InterPro; IPR000743; GH28.
RN      Pfam; PF00295; Glyco_hydro_28; 1.

```

DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 KW Hydrolyase, Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
 KM Amylolyase; Glycoprotein; Allergen.
 FT SIGNAL 1 22
 FT PROPEP 23 45
 FT CHAIN 46 433
 FT PROPEP 434 514
 FT ACT SITE 278 278
 FT CARBOHYD 460 460
 FT CARBOHYD 472 472
 FT CONFLICT 5 5
 FT CONFLICT 12 12
 FT CONFLICT 34 35
 FT CONFLICT 37 37
 FT CONFLICT 88 88
 FT CONFLICT 98 98
 FT CONFLICT 451 451
 FT CONFLICT 454 454
 FT CONFLICT 504 504
 FT CONFLICT 507 507
 SQ SEQUENCE 514 AA; 56645 MW; 62461133FAD6302 CRC64;
 Query Match 100.0%; Score 67; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.00012; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;
 Oy 1 LSDISLKTSGKIAS 15
 Db 395 LSDISLKTSGKIAS 409

RESULT 2
 ZNUC_ECOLI STANDARD; PRT; 251 AA.
 ID ZNUC_ECOLI
 AC P52648; P76285;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE High-affinity zinc uptake system ATP-binding protein znuC.
 GN ZNUC OR B1858 OR Z2910 OR EC82568.
 OS Escherichia coli, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_Taxid=562, 83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / EMG2;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251358; PubMed=9097040;
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nampak G., Seki Y., Sivasubraman S., Tagami H.,
 RA Saito N., Sampei G., Seki Y., Sivasubraman S., Tagami H.,
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horichi T.,
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 40.1-50.0 min region on the linkage map."
 RL DNA Res. 3:379-392 (1996).

[4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RC MEDLINE=21074935; PubMed=11206551;
 RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimantla E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blatner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533 (2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RUMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shimagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12."
 RL DNA Res. 8:11-22 (2001).
 RN [6]
 RP CHARACTERIZATION.
 RX MEDLINE=98343803; PubMed=9680209;
 RA Patzer S.I., Hantke K.;
 RT "The znuABC high-affinity zinc uptake system and its regulator Zur in
 Escherichia coli."
 RL Mol. Microbiol. 28:1199-1210 (1998).
 CC -1- FUNCTION: INVOLVED IN THE HIGH-AFFINITY ZINC UPTAKE TRANSPORT
 SYSTEM.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 FRAMESHIFT IN POSITION 188.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC CC
 CC EMBL; U38702; -; NOT ANNOTATED_CDS.
 DR EMBL; AE000280; AAC74928.1; -;
 DR EMBL; D90828; BAA15666.1; -;
 DR EMBL; D90829; BAA15669.1; -;
 DR EMBL; AE005408; AAG56848.1; -;
 DR EMBL; AP002559; BAB35991.1; -;
 DR Ecogene; EG13132; znuC.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transportr.
 DR Pfam; PF00005; ABC_tran.1.
 DR ProDom; PD000006; ABC_transportr.1.
 DR SMART; SM00382; AAA.1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; FALSE NEG.
 KW ATP-binding; Transport; Zinc; Complete proteome.
 FT NP_BIND 37 44
 FT CONFLICT 19 19
 FT CONFLICT 80 80
 FT CONFLICT 83 83
 FT CONFLICT 135 135
 FT CONFLICT 143 143
 FT CONFLICT 155 155
 FT CONFLICT 174 174
 FT CONFLICT 176 176
 SQ SEQUENCE 251 AA; 27867 MW; F4B845AC3C7904C CRC64;
 Query Match 64.2%; Score 43; DB 1; Length 251;
 Best Local Similarity 69.2%; Pred. No. 1.4;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSDISLKTSGK1 13
 DB 20 LSDVSLKPKGKI 32

RESULT 3

FECE_ECOLI STANDARD; PRT; 255 AA.
 AC P15031;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Iron(III) dicitrate transport ATP-binding protein fece.
 GN FECE OR P4287.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=89213950; PubMed=2651410;
 RA Staudemater H., van Hove B., Yargochi Z., Braun V.;
 RT "Nucleotide sequences of the fecBDE genes and locations of the
 RT proteins suggest a periplasmic-binding-protein-dependent transport
 RT mechanism for iron(III) dicitrate in Escherichia coli.";
 RL J. Bacteriol. 171:2626-2633(1989).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=95334362; PubMed=7610040;
 RA Bauriedl V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 RA Blatter F.R.;
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
 RT region from 92.8 through 100 minutes.";
 RL Nucleic Acids Res. 23:2105-2119(1995).
 -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
 FOR CITRATE-DEPENDENT IRON(III). PROBABLY RESPONSIBLE FOR ENERGY
 COUPLING TO THE TRANSPORT SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M28397; AAA23765.1; -;
 DR EMBL; U14003; AAA97183.1; -;
 DR EMBL; AE000499; AAC77243.1; -;
 DR PIR; J50115; QRECM3.
 DR EcoGene; EG10290; fece.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_Transport.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transport; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW Iron transport; Transport; Inner membrane; ATP-binding;
 FT Complete proteome.
 NP BIND 35
 SQ SEQUENCE 255 AA; 28190 MW; 89785C2A91D51FF3 CRC64;

Query Match 62.7%; Score 42; DB 1; Length 255;
 Best Local Similarity 53.3%; Pred. No. 2.2;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 LSDISLKTSGK1AS 15
 DB 18 LNDVSLSLPKGKITA 32

RESULT 4

YP12_AGRU STANDARD; PRT; 145 AA.
 ID YP12_AGRU
 AC 044433;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 15.8 kDa protein in plnF2.3' region (ORF2).
 OS Agrobacterium tumefaciens.
 OG Plasmid pTiA6.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=358;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89213933; PubMed=2708311;
 RA Kanemoto R.H., Powell A.T., Akiyoshi D.E., Regier D.A.,
 RA Kersteeter R.A., Nester E.W., Hawes M.C., Gordon M.P.;
 RT "Nucleotide sequence and analysis of the plant-inducible locus plnF
 RT from Agrobacterium tumefaciens.";
 RL J. Bacteriol. 171:2506-2512(1989).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M19352; AAA82504.1; -;
 DR KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 145 AA; 15829 MW; 95A3D3D8E9560A6 CRC64;

Query Match 56.7%; Score 38; DB 1; Length 145;
 Best Local Similarity 66.7%; Pred. No. 6.5;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSDISLKTSGK 12

DB 29 LSEIELALYAGK 40

RESULT 5

Y248_CAEEL STANDARD; PRT; 316 AA.
 ID Y248_CAEEL
 AC 076387;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Hypothetical 33.7 kDa protein C24G6.8 in chromosome V.
 GN C24G6.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Greco T., Bradshaw H., Keppeler D.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 -1- SIMILARITY: BELONGS TO THE UPF0099 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

GN YTRF
 OS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CX NCBI_TextID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98048467; PubMed=9387221;
 RA MEDLINE=98048467; PubMed=9387221;
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
 RT "Sequencing and functional annotation of the Bacillus subtilis genes
 in the 200 kb rnm-dnaB region";
 RL Microbiology 143:3431-3441(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Oggessara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Polochin A., Borchert S.,
 RA Bottaris R., Boursier L., Brans A., Braun M., Bridgell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Enriac K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
 RA Gilm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Hatch J., Hallio M.F., Itaya M., Jones L.,
 RA Hilbert H., Holsappel S., Hosono S., Klier-Biancard M., Klein C.,
 RA Joris B., Katamata D., Kasahara Y., Kleier-Biancard M., Klein C.,
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,
 RA Preece E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadei Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowka A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccon E., Takagi T., Takahashi H., Takemaru K.,
 RA Takuchi M., Tanakoshi A., Tanaka T., Terstra P., Tognoni A.,
 RA Toesto V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viati A., Wambut R., Wedler B., Wedler B., Weitzneger T.,
 RA Winiers P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumsstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis".
 RL Nature 390:249-256(1997).
 RL Nature 390:249-256(1997).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF008220; AAC00251.1; -
 DR EMBL: Z99119; CAB15020.1; -
 DR HSSP: 058663; 1G6H.
 DR Subtilist, BG13915; ytrF.
 DR InterPro: IPR003439; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PDOM0006; ABC_transportr; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KM Hypothetical protein_ATP-binding; (PROMOTAI);
 NT NE_BIND 42 49 ATP_BIND
 SQ SEQUENCE 231 AA; 25460 MW; FC342312BCC1172 CRC64;
 Query Match 55.2%; Score 37; DB 1; Length 231;
 Best Local Similarity 50.0%; Pred. No. 16;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 LSDISLKTSGKIA 14
 DB 25 LKOVSLVAKGEIA 38

RESULT 8

YHCG_BACSU STANDARD; PRT; 232 AA.

DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical ABC transporter ATP-binding protein yhcg.
 GN YHCG.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]

SEQUENCE FROM N.A.

RA MEDLINE=97124185; PubMed=8969498;
 RA Noback M.A., Terpetra P., Holasappel S., Venema G., Bron S.;
 RT "A 22 kb DNA sequence in the *cspB-glpPFXD* region at 75 degrees on the
 RT Bacillus subtilis chromosome."
 RL Microbiology 142:3021-3026(1996).
 RN [2]

SEQUENCE FROM N.A.

RA STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell I.F., Cummings N.J., Daniel R.A.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Danizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrati E., Rouger D.,
 RA Filiz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gilm S.Y., Glaeser P., Goiteau A., Goldightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holasappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Kleaer-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Pressan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiuchi J., Sekowska A., Setor S.J., Setor P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi R., Takemaru K.,
 RA Sotokuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vaasacotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzmeger T.,
 RA Winers P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*."
 RT Nucleic 390:249-256(1997).

-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X36983; CAA5690.1; -.

DR EMBL: Z99108; CAB12735.1; -.
 DR Subtilist; BG1585; yhcg.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; FALSE_NEG.
 DR Hypothetical protein; ATP-binding; transport; Complete proteome.
 KW NP_BIND
 FT 35
 SQ SEQUENCE 232 AA; 26520 MW; 8DACDAEBC7414FF5 CRC64;

Query Match 55.2%; Score 37; DB 1; Length 232;
 Best local similarity 46.2%; Pred. No. 16;
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSDISLKTSGK 13
 DB 18 VNDVSLTSGRI 30

RESULT 9

Y467_MYCGE STANDARD; PRT; 311 AA.

DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical ABC transporter ATP-binding protein MG467.
 GN MG467.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]

SEQUENCE FROM N.A.

RA STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Dougherty B.A., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Luetter T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of *Mycoplasma genitalium*."
 RL Science 270:397-403(1995).
 RN [2]

SEQUENCE OF 14-117 FROM N.A.

RA STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bort K.F., Hutchison C.A. III;
 RT "A survey of the *Mycoplasma genitalium* genome by using random
 RT sequencing."
 RL J. Bacteriol. 175:7918-7930(1993).

-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U39728; AAC72487.1; -.
 DR EMBL: U01741; AAD10551.1; -.

DR TIGR; MG467.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR Hypothetical protein; ATP-binding; transport; Complete proteome.
 KW

DR EMBL; M73329; AAA39944.1; -
 DR HSP; P07237; IMK.
 DR SWISS-2DPAGE; P27773; MOUSE.
 DR WED; MG1:95834; Gp58.
 DR InterPro; IPR000063; ThioRed.
 DR Pfam; PF00085; thioRed; 2.
 DR PRINTS; PR00421; THIOREDOXIN.
 DR TIGRFAMs; TIGR01126; pdi_dom; 2.
 DR TIGRFAMs; TIGR01130; ER_pdi_fam; 1.
 DR PROSITE; PS00194; THIOREDOXIN; 2.
 DR Redox-active center; Isomerase; Endoplasmic reticulum; Repeat; Signal;
 KW Phosphorylation.
 FT SIGNAL 1 24
 FT CHAIN 25 504 PROTEIN DISULFIDE ISOMERASE A3.
 FT DISULFID 57 60 REDOX-ACTIVE (BY SIMILARITY).
 FT DISULFID 405 408 REDOX-ACTIVE (BY SIMILARITY).
 FT SITE 501 504 PREVENT SECRETION FROM ER
 (BY SIMILARITY).
 SQ SEQUENCE 504 AA; 56621 MW; 5983B04755CEB7B4 CRC64;
 Query Match 55.2%; Score 37; DB 1; Length 504;
 Best Local Similarity 53.8%; Pred. No. 38;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSDISLKLTSKGI 13
 DB 310 LSDISLKLTSKGI 322
 RESULT 12
 CG18 YEAST STANDARD; PRT; 229 AA.
 ID CG18 YEAST STANDARD; PRT; 229 AA.
 AC P38754;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GI/S-specific cyclin PCL5.
 GN PCL5 OR YHR071W
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 ON NCBI_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8286C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaastis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 VII.",
 RT Science 265:2077-2082(1994).
 CC -1- SIMILARITY: BELOWS TO THE CYCLIN FAMILY. MOST SIMILAR TO GI/S-
 SPECIFIC CYCLIN PCL1 AND PCL2.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; U00061; AAB68375.1; -
 DR PIR; S46696; S46696.
 DR SGD; S0001113; PCL5.
 DR InterPro; IPR004366; Cyclin.
 DR Pfam; PF00134; cyclin; 1.
 DR SMART; SM00385; CYCLIN; 1.

DR PROSITE; PS00292; CYCLIN; FALSE NEG.
 KW Cyclin; Cell cycle; Cell division.
 SQ SEQUENCE 229 AA; 26467 MW; B6839AB9D85DD49 CRC64;
 Query Match 53.7%; Score 36; DB 1; Length 229;
 Best Local Similarity 61.5%; Pred. No. 25;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSDISLKLTSKGI 13
 DB 61 LSDISLKLTSKGI 73
 RESULT 13
 MNTB BACSU STANDARD; PRT; 250 AA.
 ID MNTB BACSU STANDARD; PRT; 250 AA.
 AC Q34338;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Manganese transport system ATP-binding protein mntB.
 GN MNTB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 ON NCBI_TaxID=1423;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell E., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Dentzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Eutlian K.D., Eyrington J., Fabret C., Ferrati E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghm S.Y., Glaser P., Goffeau A., Colighly E.J., Grandi G.,
 RA Guisepi G., Guy B.U., Haga K., Haeck J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Ilaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianhard M., Klein C.,
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogilwa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Priescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sada Y.,
 RA Sato T., Scanlan E., Schleich S., Schoefer R., Scoffone P.,
 RA Sekiguchi J., Sekowska A., Serr S.J., Serr P., Shin B.S., Soldo B.,
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Toosto V., Uchiyama S., Vandenbol M., Vannier F., Vasserolet A.,
 RA Viati A., Wambut R., Wedler E., Wedler H., Welzenegger T.,
 RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasunoto K., Yeta K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RT Nature 390:249-256(1997).
 [3]
 POSSIBLE FUNCTION.
 RP MEDLINE=20223631; PubMed=10760146;
 RX Que O., Helmann J.D.;
 RA "Manganese homeostasis in Bacillus subtilis is regulated by MntR, a

```

RT bifunctional regulator related to the diphtheria toxin repressor
RT family of proteins."
RL Mol. Microbiol. 35:1454-1468(2000).
CC -1- FUNCTION: This protein is probably a component of a manganese
CC perennase, a binding protein-dependent, ATP-driven transport
CC system. Probably responsible for energy coupling to the transport
CC system.
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF008220; AAC00230.1; -.
DR EMBL, Z99119; CAB15054.1; -.
DR Subtilist; BG13852; mntb.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00005; ABC_tran.1.
DR Prodom; PD000006; ABC_transportr; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR Transport; ATP-binding; Membrane; Complete proteome.
DR NP BIND 36 43 ATP (POTENTIAL).
SQ SEQUENCE 250 AA; 27881 MW; 12A386749E2208B CRC64;

Query Match 53.7%; Score 36; DB 1; Length 250;
Best Local Similarity 53.8%; Pred. No. 27;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSDISLKLTSGLK 13
DB 19 LSDISLQVPEGLK 31

RESULT 14
CYSB_ECOLI STANDARD; PRT; 365 AA.
AC P16676; P76993;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sulfate transport ATP-binding protein cysA.
GN CYSB OR B2422.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
ON [1]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RC MEDLINE=90264334; PubMed=2189958;
RA Siro A., Hymelwitz M.M., Hulanicka D.M., Boeck A.,
RA "Sulfate and thiosulfate transport in Escherichia coli K-12:
RT nucleotide sequence and expression of the cystWAM gene cluster.";
RL J. Bacteriol. 172:3351-3357(1990).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / W61655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.,
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RN SEQUENCE, FROM N.A.

```

```

RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayaishi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sampaio G., Satoh Y., Sivaraman S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features."
RL DNA Res. 4:91-113(1997).
CC -1- FUNCTION: INVOLVED IN THE HIGH-AFFINITY SULFATE AND THIOSULFATE
CC MEMBRANE TRANSPORT SYSTEM.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M32101; AAA2639.1; -.
DR EMBL, AB000329; AAC75475.1; -.
DR EMBL, D90872; BAA16305.1; -.
DR EMBL, D90871; BAA16296.1; -.
DR PIR; C35402; QRECSA.
DR EcoGene; EG10183; cysA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00005; ABC_tran.1.
DR Prodom; PD000006; ABC_transportr; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00968; 3a0106s01.1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR Sulfate transport; Transport; ATP-binding; Inner membrane;
KM Complete proteome.
FT NP BIND 35 42 ATP (BY SIMILARITY).
FT CONFLICT 136 137 QL -> HV (IN REF. 1).
SQ SEQUENCE 365 AA; 41059 MW; B5FCC346EDF2788 CRC64;

Query Match 53.7%; Score 36; DB 1; Length 365;
Best Local Similarity 46.7%; Pred. No. 41;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSDISLKLTSGLK 15
DB 18 LSDISLQVPEGLK 32

RESULT 15
CYSB_SALTY STANDARD; PRT; 365 AA.
AC P40860;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sulfate transport ATP-binding protein cysA.
GN CYSB OR STM2441.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
ON [1]
RN SEQUENCE FROM N.A.
RP STRAIN=LT2 / SGC1412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=1677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,

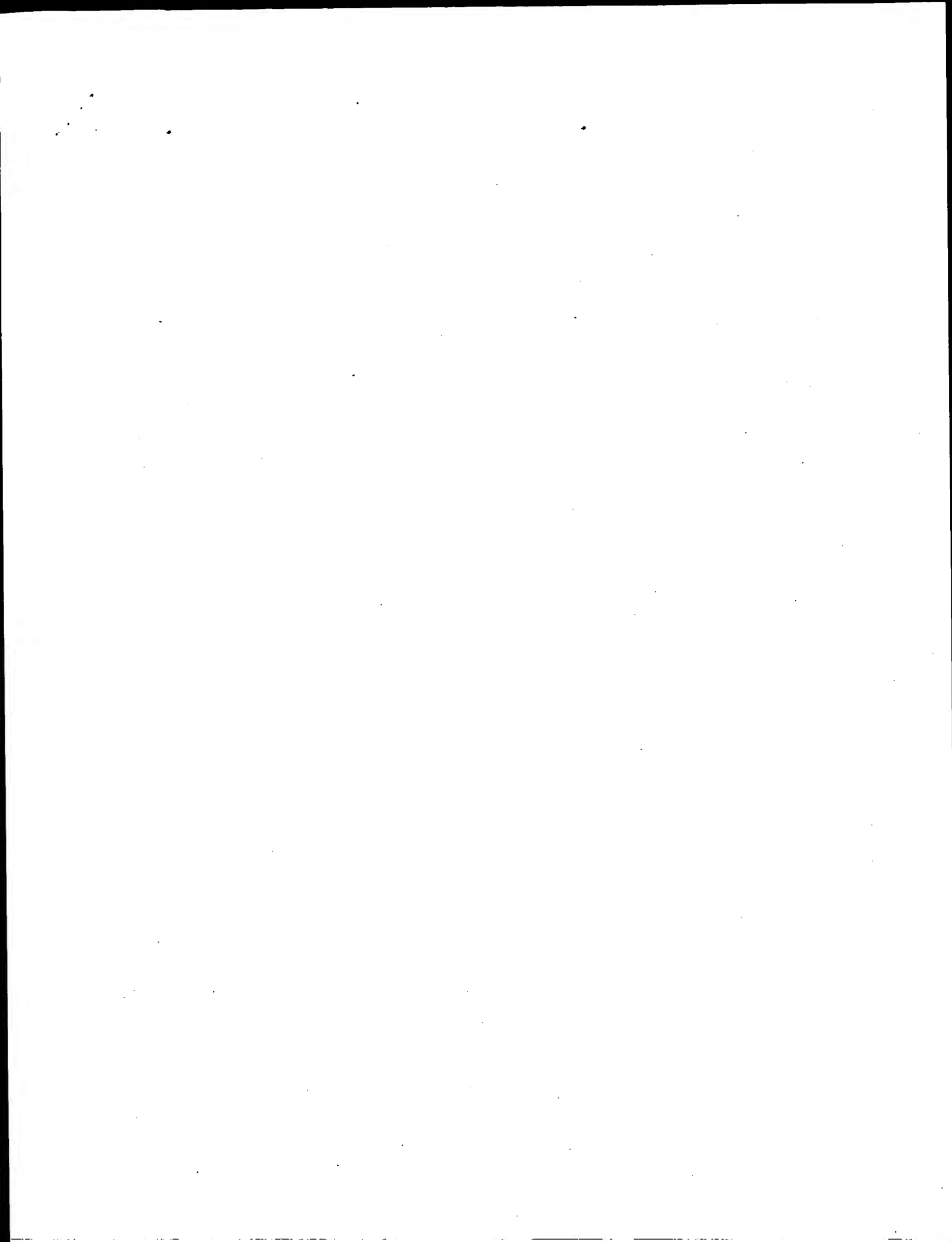
```

RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856 (2001).
 RN [2]
 RP SEQUENCE OF 248-365 FROM N.A.
 RC STRAIN=LT2;
 RA Sivaprasad A.V., Kuczek E.S., Bawden C.S., Rogers G.E.;
 RL Submitted (May-1991) to the EMBL/GenBank/DBJ databases
 CC -! FUNCTION: INVOLVED IN THE HIGH-AFFINITY SULFATE AND THIOSULFATE
 CC MEMBRANE TRANSPORT SYSTEM.
 CC -! SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
 CC -! SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB008810; AAL21335.1; -;
 DR EMBL; X59595; -; NOT_ANNOTATED_CDS.
 DR StyGene; SG10518; CYSA.
 DR InterPro; IPR003439; ABC_transportr.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transportr; 1.
 DR TIGRFAMs; TIGR00968; 3a0106601; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KM Sulfate transport; Transport; ATP-binding; Inner membrane;
 KW Complete proteome.
 FT NP_BIND 35 42 ATP (BY SIMILARITY).
 FT CONFLICT 298 303 EASPKG -> SOPS (IN REF. 2).
 SQ SEQUENCE 365 AA; 41035 MW; 0166EAC2AF2B1C4C CRC64;

Query Match 53.7%; Score 36; DB 1; Length 365;
 Best Local Similarity 46.7%; Pred. No. 41;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSDISLKLTKSGKIAS 15
 :|||:|:|:|:
 Db 18 LNDISLDIPSGQMV 32

Search completed: April 20, 2003, 13:07:47
 Job time : 4.92105 secs



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds
(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524D-28

Perfect score: 81

Sequence: 1 RPLMTIFSGNNIKL 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneeqp-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneeqp-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneeqp-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneeqp-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneeqp-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneeqp-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneeqp-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneeqp-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneeqp-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneeqp-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneeqp-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneeqp-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneeqp-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneeqp-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneeqp-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneeqp-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneeqp-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneeqp-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneeqp-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneeqp-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneeqp-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneeqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	20	15	AA45548
2	81	100.0	20	16	AA45548
3	81	100.0	30	19	AA45548
4	81	100.0	134	18	AA45548
5	81	100.0	353	16	AA45548
6	81	100.0	353	17	AA45548
7	81	100.0	373	20	AA45548
8	81	100.0	373	20	AA45548
9	81	100.0	374	14	AA45548
10	81	100.0	374	15	AA45548

11	81	100.0	374	15	AA45548
12	81	100.0	374	16	AA45548
13	81	100.0	374	20	AA45548
14	81	100.0	374	20	AA45548
15	62	76.5	354	17	AA45548
16	62	76.5	354	17	AA45548
17	62	76.5	354	17	AA45548
18	62	76.5	354	17	AA45548
19	62	76.5	354	17	AA45548
20	62	76.5	354	17	AA45548
21	57	70.4	501	22	AA45548
22	57	70.4	501	22	AA45548
23	57	70.4	501	22	AA45548
24	57	70.4	501	22	AA45548
25	57	70.4	501	22	AA45548
26	56	69.1	455	23	AA45548
27	56	69.1	455	23	AA45548
28	56	69.1	455	23	AA45548
29	56	69.1	455	23	AA45548
30	56	69.1	455	23	AA45548
31	56	69.1	455	23	AA45548
32	56	69.1	455	23	AA45548
33	56	69.1	455	23	AA45548
34	56	69.1	455	23	AA45548
35	56	69.1	455	23	AA45548
36	56	69.1	455	23	AA45548
37	56	69.1	455	23	AA45548
38	56	69.1	455	23	AA45548
39	56	69.1	455	23	AA45548
40	56	69.1	455	23	AA45548
41	56	69.1	455	23	AA45548
42	56	69.1	455	23	AA45548
43	56	69.1	455	23	AA45548
44	56	69.1	455	23	AA45548
45	56	69.1	455	23	AA45548

ALIGNMENTS

RESULT 1
AA45548
ID AA45548 standard; Protein; 20 AA.
XX
AC AA45548;
XX
DT 13-JUL-1994 (first entry)
XX
DE Cry j I pollen allergen peptide CRI-7.
XX
KW Japanese cedar; detection; allergy; treatment; diagnosis;
KW T cell epitope; sensitivity.
XX
OS Cryptomeria japonica.
XX
PN WO9401560-A.
XX
PD 20-JAN-1994.
XX
PF 15-FAN-1993; 93WO-US00139.
XX
PR 01-SEP-1992; 92US-0938990.
XX
PR 10-JUL-1992; 92WO-US05661.
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PI Bond UF, Garman RD, Griffith IU, Kuo M, Pollock J;
XX WPI, 1994-03506/04.
DR
XX Antigen derived from Japanese cedar pollen allergen Cry j I -
XX contain at least two T cell epitope(s), used to treat or diagnose
XX allergy

Japanese cedar pol
Cry j I Japanese C
cedar allergen 493
Japanese cedar all
T-cell epitope pep
Chamaecyparis obtu
Chamaecyparis cypres p
Jun s I. Juniperu
Chamaecyparis obtu
Herbicidally activ
Japanese cedar pol
Residues 61-75 of
Amino acid sequenc
Herbicidally activ
Herbicidally activ
Herbicidally activ
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Part of Amb a I/An
Ragweed allergen U
UNC Clone 6 Amb a
Amb a I/Antigen E
Ragweed Pollen All
Ragweed Amb a I.2 a
Allergen Amb a I.8
Ragweed allergen 1
Full length Amb a
Full length Amb a
Ragweed pollen Amb
Ragweed pollen UNC

XX Claim 1; Fig 13; 137bp; English.

PS The sequence is that of an isolated peptide of the Japanese cedar
CC pollen allergen Cry j I (amino acids 61-80). The peptide, CJI-7,
CC can be used for the treatment and diagnosis of allergies associated
CC with Japanese cedar pollen. It has enhanced therapeutic properties
CC but reduced side effects compared to naturally occurring allergens.
XX

Sequence 20 AA;

Query Match 100.0%; Score 81; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPLWIFSGNNIKL 15
DB 6 RPLWIFSGNNIKL 20

RESULT 2

AA82497
ID AAR82497 standard; Protein; 20 AA.

AC AAR82497;

DT 15-APR-1996 (first entry)

DE Cry j I Japanese Cedar pollen allergen peptide fragment (CJI-7).

KM Cry j I; Japanese cedar pollen allergen; modified; drug production;

KW allergy; Cryptomeria japonica.

XX Cryptomeria japonica.

OS WO9227786-A1.

PN 19-OCT-1995.

PF 06-APR-1995; 95WO-US04249.

PR 06-DEC-1994; 94US-0350225.

PR 08-APR-1994; 94US-0226248.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

XX Chen X, Evans S, Franzen HM, Kuo M, Powers SP;

PI Shaked Z;

DR WPI; 1995-366391/47.

XX Modified Cryptomeria japonica (Cry j) I peptide(s) - useful for
PT treating allergy to Japanese cedar pollen allergen or
PT immunologically cross reactive allergens
XX

PS Disclosure; Figure 2; 60bp; English.

XX Novel peptides of cry j I have been modified as a part of a
CC preformulation scheme to develop an optimised drug product for
CC therapeutic treatment of humans suffering from allergy to Japanese
CC cedar pollen allergen or an allergen which is immunologically cross
CC reactive with Japanese cedar pollen allergen. Such modified peptides
CC possess certain characteristics which render them particularly
CC suitable for drug product formulation. Peptide fragments of Cry j I,
CC modified and unmodified, are given in AAR82491-R82525. This peptide
CC fragment corresponds to amino acids 61-80 of the allergen mature
CC protein.
XX

Sequence 20 AA;

Query Match 100.0%; Score 81; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPLWIFSGNNIKL 15
DB 6 RPLWIFSGNNIKL 20

RESULT 3

AAW44682
ID AAW44682 standard; peptide; 30 AA.

AC AAW44682;

DT 01-MAY-1998 (first entry)

DE T-cell epitope peptide #1 of sugi pollen antigen.

KM T-cell epitope; sugi pollen antigen; sugi pollinosis.

XX Synthetic.

OS Cryptomeria japonica.

PN JP10007700-A.

PD 13-JAN-1998.

PF 24-JUN-1996; 96JP-0163287.

PR 24-JUN-1996; 96JP-0163287.

PA (DAIL) DAICEL CHEM IND LTD.

PA (MEIT) MEIT SEIKA KAISHA LTD.

DR WPI; 1998-133630/13.

XX T cell epitope peptide of sugi pollen antigen - useful in the

PT treatment of sugi pollinosis

XX Claim 1; Page 1; 14pp; Japanese.

CC T-cell epitope peptides AAW44682-88 and their derivatives react with
CC sugi pollinosis patient peripheral blood T lymphocytes. A composition
CC prepared by combining at least 2 of the above peptides and/or their
CC derivatives is used for the prevention and treatment of sugi
CC pollinosis.
XX

Sequence 30 AA;

Query Match 100.0%; Score 81; DB 19; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPLWIFSGNNIKL 15
DB 6 RPLWIFSGNNIKL 20

RESULT 4

AAW27371
ID AAW27371 standard; peptide; 134 AA.

AC AAW27371;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #3.

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;

KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN WO9732600-A1.

XX

	RESULT 5
ID	AAR75388
XX	AAR75388 standard; protein; 353 AA.
AC	
XX	AAR75388;
DT	12-MAR-1996 (first entry)
XX	
DE	Japanese cedar pollen allergen Cryj I.
KM	Japanese cedar; pollen allergen; Cryj I; T-cell epitopes; peptides; prevention; treatment; cryptomeria pollinosis.
XX	
OS	Cryptomeria japonica.
XX	
Key	Location/Qualifiers
FH	Peptide
FT	61..75
/note=	"T-cell epitope peptide"
FT	91..105
Peptide	/note= "T-cell epitope peptide"
FT	106..120
Peptide	/note= "T-cell epitope peptide"
FT	146..160
Peptide	/note= "T-cell epitope peptide"
FT	211..225
Peptide	/note= "T-cell epitope peptide"
FT	326..340
Peptide	/note= "T-cell epitope peptide"
FT	335..346
Peptide	/note= "T-cell epitope peptide"
XX	
PN	JP07118295-A.
XX	

DT	24-MAY-1996	(first entry)
XX		
XX		
DE	Cedar pollen allergen B.	
KW	Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope; antibody; pollinosis; therapy; immunotherapy.	
KW		
XX		
OS	Cryptomeria japonica.	
PN	EP700929-A2.	
XX		
PD	13-MAR-1996.	
XX		
PF	08-SEP-1995; 95SEP-0306295.	
XX		
PR	14-JUL-1995; 95JUL-0200221.	
PR	10-SEP-1994; 94JUL-0242137.	
PR	14-JUL-1995; 95JUL-0200204.	
XX		
XX		
PA	(HAYB) HAYASHIBARA SEIBUTSU KAGAKU.	
PI	Hino K, Saito S, Taniguchi Y;	
XX		
DR	WPI; 1996-140976/15.	
PT	New peptide(s) derived from cedar pollen allergens - activate	
PT	allergen-specific T-cells, but not allergen-specific IgE antibodies,	
PT	used for treating cedar pollinosis	
PS	Claim 5; Page 31-32; 36pp; English.	
XX		
CC	Synthetic peptides based on portions of cedar pollen allergens A	
CC	(AR81566) and B (AR81567) were tested for their ability to activate	
CC	cedar allergen-specific T-cells, but not allergen-specific IgE	
CC	antibodies. 6 Peptides (AR81560-81585) were identified as T-cell	
CC	epitopes. These peptides, plus subsequences (AR81573-79) essential	

CC for T-cell recognition, and homologous peptides (AA081588-96) can
CC be used as immunotherapeutic agents to treat or prevent cedar
CC pollinosis, avoiding side-effects such as anaphylaxis.

SO Sequence 353 AA.

Query Match 100.0%; Score 81; DB 17; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 RPLWIFSGNNMNTKL 15
Db 66 RPLWIFSGNNMNTKL 80

RESULT 7
ID AAY25664
AAV25664 standard; protein; 373 AA.

AC AAY25664;

DT 30-SEP-1999 (first entry)

DE Cedar allergen 493634 Cry j IB protein fragment.

XX Major histocompatibility complex; class II; desensitizing; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.

XX WO9334826-A1.

XX 15-JUL-1999.

XX 11-JAN-1999; 99WO-GB00080.

XX 21-SEP-1998; 98GB-0020474.

XX 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Kay AB, Larche M;

XX WPI; 1999-458255/38.

XX Desensitizing patients to polypeptide allergens

XX Example 6; Page 73; 117pp; English.

XX This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a cedar (Cedrus sp.) allergen 493634 Cry j IB.

XX Sequence 373 AA;

Query Match, 100.0%; Score 81; DB 20; Length 373;

Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPLWIFSGNNMNTKL 15
Db 86 RPLWIFSGNNMNTKL 100

RESULT 8
ID AAY25668
AAV25668 standard; protein; 373 AA.

AC AAY25668;

DT 30-SEP-1999 (first entry)

DE Japanese cedar allergen 541803 Cry j I precursor protein fragment.

XX Major histocompatibility complex; class II; desensitizing; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.

XX WO9334826-A1.

XX 15-JUL-1999.

XX 11-JAN-1999; 99WO-GB00080.

XX 21-SEP-1998; 98GB-0020474.

XX 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Kay AB, Larche M;

XX WPI; 1999-458255/38.

XX Desensitizing patients to polypeptide allergens

XX Example 6; Page 75; 117pp; English.

XX This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a Japanese cedar (Cedrus sp.) allergen 541803 Cry j I
CC precursor.

XX Sequence 373 AA;

Query Match 100.0%; Score 81; DB 20; Length 373;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPLWIFSGNNMNTKL 15

Db 86 RPLWIFSGNNMNTKL 100


```

RESULT 9
ID AAR31937 standard; Protein; 374 AA.
XX
XX AAR31937;
AC
XX 03-JUN-1993 (first entry)
XX
XX Cry j I.
XX
XX Japanese cedar pollen; allergen; antigen; allergy; B cell; T cell.
XX
XX Cryptomeria japonica.
OS
XX
XX Key Location/Qualifiers
XX FT Peptide 1..21
XX FT /note= "signal peptide"
XX FT Protein 22..374
XX FT /note= "mature Cry j I"
XX
XX WO9301213-A.
XX
XX 21-JAN-1993.
XX
XX 10-JUL-1992; 92WO-US05661.
XX
XX 12-JUL-1991; 91US-0729134.
XX
XX 15-JUL-1991; 91US-0730452.
XX
XX (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
XX Bond JF, Griffith J, Pollock J;
XX
XX WPI; 1993-045434/05.
XX
XX N-PSDB; AAQ35304.
XX
XX Nucleic acid sequence encoding Cryptomeria japonica allergen -
XX PT for the diagnosis treatment and prevention of allergic reactions
XX PT to Japanese cedar pollen
XX
XX Claim 11; Page 42; 69pp; English.
XX
XX Fresh pollen and staminate cone samples were collected from a single
XX Cryptomeria japonica (Japanese cedar) tree. RNA was prep'd. and used
XX to synthesize cDNA. The cDNA was subjected to successive rounds of
XX PCR to yield a full length Cry j I clone. Cry j I or an antigenic
XX fragment of it may be used for detecting, treating and preventing an
XX allergic response to Japanese cedar pollen allergen. It is capable of
XX modifying both the B and T cell response to Cry j I and T cell response
XX to a Cry j I antigen.
XX
XX Sequence 374 AA;
SQ
Query Match 100.0%; Score 81; DB 14; Length 374;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPLMIFSGNMNIXL 15
DB 87 RPLMIFSGNMNIXL 101
RESULT 10
ID AAR45541 standard; Protein; 374 AA.
XX
XX AAR45541;
AC
XX 13-JUL-1994 (first entry)
XX
XX Cry j I pollen allergen.
XX
XX

```

```

XX
XX Japanese cedar; detection; allergy; treatment; diagnosis;
XX KM T cell epitope; sensitivity.
XX
XX Cryptomeria japonica.
OS
XX
XX Key Location/Qualifiers
XX FT Peptide 1..21
XX FT /note= "signal peptide"
XX FT Peptide 22..374
XX FT /note= "mature peptide"
XX
XX WO9401560-A.
XX
XX 20-JAN-1994.
XX
XX 15-JAN-1993; 93WO-US00139.
XX
XX 01-SEP-1992; 92US-0938990.
XX
XX 10-JUL-1992; 92WO-US05661.
XX
XX (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
XX Bond JF, Garman RD, Griffith J, Kuo M, Pollock J;
XX
XX WPI; 1994-035066/04.
XX
XX N-PSDB; AAQ55271.
XX
XX Antigen derived from Japanese cedar pollen allergen Cry j I -
XX PT contain at least two T cell epitope(s), used to treat or diagnose
XX PT allergy
XX
XX Disclosure; Fig 4; 137pp; English.
XX
XX The sequence is that of the Japanese cedar pollen allergen
XX Cry j I which contains at least two T cell epitopes. Peptide
XX antigens derived from it can be used for the treatment and
XX diagnosis of allergies associated with Japanese cedar pollen.
XX CC The peptides have enhanced therapeutic properties but reduced
XX side effects compared to naturally occurring allergens.
XX
XX Sequence 374 AA;
SQ
Query Match 100.0%; Score 81; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RPLMIFSGNMNIXL 15
DB 87 RPLMIFSGNMNIXL 101
RESULT 11
ID AAR60166 standard; Protein; 374 AA.
XX
XX AAR60166;
AC
XX 24-MAR-1995 (first entry)
XX
XX Japanese cedar pollen antigen CryjI.
XX
XX Japanese cedar; pollen antigen; allergen; CryjI; sugi; pollinosis.
XX
XX Cryptomeria japonica.
OS
XX
XX Key Location/Qualifiers
XX FT Peptide 1..21
XX FT /label= signal_peptide
XX FT Protein 22..374
XX FT /label= mature_CryjI
XX
XX JP06197768-A.
XX

```

XX 19-JUL-1994.
 PD 07-JAN-1993; 93JP-0001116.
 XX 07-JAN-1993; 93JP-0001116.
 XX (MEIJ) MEIJI SEIKA KAISHA.
 PA WPI; 1994-268680/33.
 DR N-PSDB; AAQ71601.
 XX
 XX
 PT Sugi (Japanese cedar) pollen antigen CryjI - is useful for
 PT diagnosis, treatment and prevention of sugi pollinosis
 XX
 PS Claim 2; Page 5-7, 9pp; Japanese.
 XX
 CC The coding sequence for the Japanese cedar ("sugi") pollen allergen
 CC CryjI was isolated from a cDNA library prepared from polyA mRNA. All
 CC or part of the CryjI protein can be used for diagnosis, treatment
 CC and prevention of sugi pollinosis.
 XX
 SQ Sequence 374 AA;
 Query Match 100.0%; Score 81; DB 15; Length 374;
 Best Local Similarity 100.0%; Pred. No. 3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RPLWIFSGNNIKL 15
 |||||
 DB 87 RPLWIFSGNNIKL 101
 RESULT 12
 AAR82490
 ID AAR82490 standard; Protein; 374 AA.
 XX
 AC AAR82490;
 XX
 DT 15-APR-1996 (first entry)
 XX
 DE Cry j I Japanese Cedar pollen allergen.
 XX
 DE Cry j I Japanese cedar pollen allergen; modified; drug production;
 KM allergy; Crytpomeria japonica.
 XX
 OS Crytpomeria japonica.
 XX
 PN WO9527786-A1.
 XX
 PD 19-OCT-1995.
 XX
 PF 06-APR-1995; 95WO-US04249.
 XX
 PR 06-DEC-1994; 94US-0350225.
 PR 08-APR-1994; 94US-0226248.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP;
 PI Shaked Z;
 DR WPI, 1995-366391/47.
 DR N-PSDB; AAT04248.
 XX
 XX Modified Crytpomeria japonica (Cry j) I peptide(s) - useful for
 PT treating allergy to Japanese cedar pollen allergen or
 PT immunologically cross reactive allergens
 XX
 PS Disclosure; Figure 1; 60pp; English.
 CC Novel peptides of cry j I have been modified as a part of a
 CC preformulation scheme to develop an optimised drug product for

CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j I,
 CC modified and unmodified, are given in AAR82491-R02525.
 XX
 SQ Sequence 374 AA;
 Query Match 100.0%; Score 81; DB 16; Length 374;
 Best Local Similarity 100.0%; Pred. No. 3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RPLWIFSGNNIKL 15
 |||||
 DB 87 RPLWIFSGNNIKL 101
 RESULT 13
 AAY25665
 ID AAY25665 standard; protein; 374 AA.
 XX
 AC AAY25665;
 XX
 DT 30-SEP-1999 (first entry)
 XX
 DE Cedar allergen 493632 Cry j IA protein fragment.
 XX
 DE Major histocompatibility complex; class II; desensitizing; human;
 KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX
 XX Cedrus sp.
 OS
 XX
 XX WO9934826-A1.
 XX
 PN 15-JUL-1999.
 XX
 PD 11-JAN-1999; 99WO-GB00080.
 XX
 PF 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Kay AB, Larche M;
 PI WPI, 1999-458255/38.
 XX
 DR Desensitizing patients to polypeptide allergens
 XX
 PS Example 6; Page 73; 117pp; English.
 XX
 CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, silkworm, honeybee,
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, larvae of
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a cedar (Cedrus sp.) allergen 493632 Cry j IA.

SQ Sequence 374 AA;
 Query Match 100.0%; Score 81; DB 20; Length 374;
 Best Local Similarity 100.0%; Pred. No. 3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPLWIFSGNNIKL 15
 |||||
 DB 87 RPLWIFSGNNIKL 101

RESULT 14

AAV25669
 ID AAV25669 standard; protein; 374 AA.

AC AAV25669;

DT 30-SEP-1999 (first entry)

DE Japanese cedar allergen 541802 Cry j I precursor protein fragment.

KM Major histocompatibility complex; class II; desensitizing; human;

KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;

KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;

KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;

KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;

KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.

PN WO9934826-A1.

PD 15-JUL-1999.

PF 11-JAN-1999; 99WO-GB00080.

PR 21-SEP-1998; 98GB-0020474.

PR 09-JAN-1998; 98GB-0000445.

PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

PI Kay AB, Larche M;

PT WPI; 1999-458255/38.

PS Desensitizing patients to polypeptide allergens

XX Example 6; Page 75; 117pp; English.

CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC class II molecule. The methods can be used for
 CC desensitising patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Teniprio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 541802 Cry j I
 CC precursor.

SQ Sequence 374 AA;

Query Match 100.0%; Score 81; DB 20; Length 374;
 Best Local Similarity 100.0%; Pred. No. 3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPLWIFSGNNIKL 15
 |||||
 DB 87 RPLWIFSGNNIKL 101

RESULT 15

AAW42129
 ID AAW42129 standard; peptide; 20 AA.

AC AAW42129;

DT 16-JUN-1998 (first entry)

DE T-cell epitope peptide 9 from Japanese cypress pollen antigen Chaol.

KW Japanese cypress pollen; antigen; T-cell epitope; Chaol; Chaol2;

KW diagnosis; allergy; spring tree pollen disease; pollinosis.

OS Chamaecyparis obtusa.

PN WO9747648-A1.

PD 18-DEC-1997.

PF 12-JUN-1997; 97WO-JP02031.

PR 14-JUN-1996; 96JP-0153527.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Kuno K;

PD WPI; 1998-052242/05.

PT T-cell epitope peptide portion of Japanese cypress pollen antigens

PT Chaol and Chaol2 - used for diagnosis and treatment of spring tree

PS pollen disease

PS Claim 1; Page 21; 71pp; Japanese.

CC The present sequence represents a T-cell epitope peptide from Japanese

CC cypress pollen antigen Chaol. The present invention describes peptides

CC which correspond to the T-cell epitope sites on Japanese cypress pollen

CC antigens Chaol and Chaol2. The peptides can be used as a reagent for the

CC diagnosis of allergy to Japanese cypress pollen, and as an antigen in

CC the treatment and prevention of spring tree pollen disease in which the

CC pollinosis involves reactivity to Japanese cypress pollen.

XX Search completed: April 20, 2003, 13:06:04

XX Job time : 19.1974 secs

XX

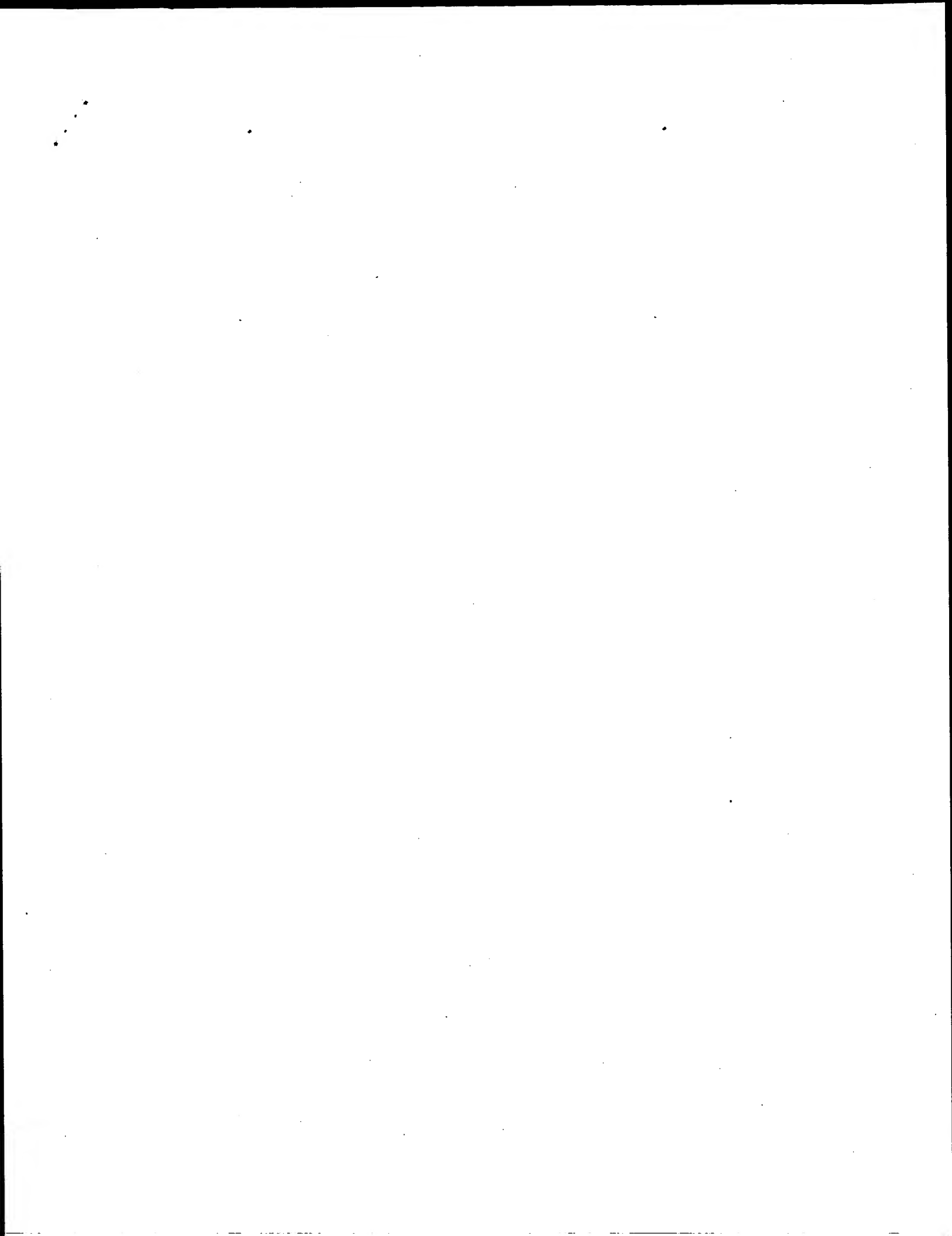
XX

XX

XX

XX

XX



GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds
(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524D-29

Perfect score: 1 IFSGNNMIXLMPY 15

Sequence: 1 IFSGNNMIXLMPY 15

Scoring table: BLAST62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/geneqp-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneqp-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneqp-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneqp-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneqp-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneqp-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneqp-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneqp-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneqp-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneqp-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneqp-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneqp-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneqp-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneqp-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneqp-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneqp-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneqp-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneqp-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneqp-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneqp-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneqp-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneqp-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	20	AA45549	Cry j I pollen all
2	79	100.0	20	AA82498	Cry j I Japanese C
3	79	100.0	30	AA44682	T-cell epitope pep
4	79	100.0	134	AA27371	Multi-epitope pep
5	79	100.0	353	AA75388	Japanese cedar pol
6	79	100.0	353	AA81587	Japanese cedar pol
7	79	100.0	373	AA25664	Cedar pollen aller
8	79	100.0	373	AA25668	Cedar pollen aller
9	79	100.0	374	AA31937	Cry j I. Cryptome
10	79	100.0	374	AA45541	Cry j I pollen all

11	79	100.0	374	AA60166	Japanese cedar pol
12	79	100.0	374	AA82490	Cry j I Japanese C
13	79	100.0	374	AA25665	Cedar allergen 493
14	79	100.0	374	AA25669	Japanese cedar all
15	68	86.1	367	AA45578	Jun s I. Juniperu
16	68	86.1	370	AA45578	Jun s I. Juniperu
17	60	75.9	20	AA42130	T-cell epitope pep
18	60	75.9	31	AA27372	Multi-epitope pep
19	60	75.9	354	AA04344	Chamaecyparis obtu
20	60	75.9	354	AA42121	Japanese cypress p
21	60	75.9	375	AAW04345	Chamaecyparis obtu
22	50	63.3	20	AA45548	Cry j I pollen all
23	50	63.3	20	AA45548	Cry j I Japanese C
24	43	54.4	42	AA07946	Human secreted pro
25	43	54.4	864	AA03636	Hypoxia-regulated
26	42	53.2	543	AA48905	Ovine adenovirus 2
27	40	50.6	104	AA004575	Human polypeptide
28	40	50.6	299	AA87761	Human T2R31 amino a
29	40	50.6	309	AAU1384	Human T2R61 (T2R6
30	40	50.6	341	AA34875	Chlamydia pneumoni
31	40	50.6	563	AAU34169	Staphylococcus aur
32	40	50.6	578	AAU36613	Staphylococcus aur
33	40	50.6	1069	AAU36613	Drosophila melanog
34	39	49.4	20	AAW42129	T-cell epitope pep
35	39	49.4	115	AAW49160	Human ubiquitin C-
36	39	49.4	257	ABG10076	Novel human diagno
37	39	49.4	391	AAU81226	Human lung cancer
38	39	49.4	402	ABG13064	Novel human diagno
39	39	49.4	532	ABG08958	Novel human diagno
40	39	49.4	532	ABG10067	Novel human diagno
41	39	49.4	1162	ABG03906	Novel human diagno
42	38	48.1	310	AA836491	Bacillus thuringie
43	38	48.1	344	ABP39330	Staphylococcus epi
44	38	48.1	371	AA611486	Arabidopsis thalia
45	38	48.1	371	AA639265	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AA45549
ID AA45549 standard; Protein; 20 AA.
XX
XX
AC AA45549;
XX
XX
DT 13-JUL-1994 (first entry)
XX
XX
DE Cry j I pollen allergen peptide CJI-8.
XX
XX
KW Japanese cedar; detection; allergy; treatment; diagnosis;
XX T cell epitope; sensitivity.
XX
XX
OS Cryptomeria-japonica.
XX
XX
PN WO9401560-A.
XX
XX
PD 20-JAN-1994.
XX
XX
PF 15-JAN-1993; 93WO-US00139.
XX
XX
PR 01-SEP-1992; 92US-0938990.
XX
XX
PR 10-JUL-1992; 92WO-US05661.
XX
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
XX
FI Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;
XX
XX
XX WPI; 1994-035066/04.
XX
XX
PT Antigens derived from Japanese cedar pollen allergen Cry j I -
PT contain at least two T cell epitope(s), used to treat or diagnose
PT allergy

XX Claim 1; Fig 13; 137bp; English.

PS The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j I (amino acids 71-90). The peptide, Cui-8,
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen. It has enhanced therapeutic properties
 CC but reduced side effects compared to naturally occurring allergens.

XX Sequence 20 AA;

Query Match 100.0%; Score 79; DB 15; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.1e-08; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFSGNNNIKLKMPY 15
 DB 1 IFSGNNNIKLKMPY 15

RESULT 2

AA82498 AAR82498 standard; Protein; 20 AA.

AC AAR82498;

DT 15-APR-1996 (first entry)

DE Cry j I Japanese Cedar pollen allergen peptide fragment (Cui-8).

KM Cry j I; Japanese cedar pollen allergen; modified; drug production;

XX allergy; Crypomeria japonica.

OS Crypomeria japonica.

PN WO9527786-A1.

PD 19-OCT-1995.

PF 06-APR-1995; 95WO-US04249.

PR 06-DEC-1994; 94US-0350225.

PR 08-APR-1994; 94US-0226248.

PA (IMMU-) IMMULOGIC PHARM CORP.

PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP;

PI Shaked Z;

PT Modified Cryptomeria japonica (Cry j) I peptide(s) - useful for

PT treating allergy to Japanese cedar pollen allergen or

PT immunologically cross reactive allergens

PS Disclosure; Figure 2; 60pp; English.

XX Novel peptides of cry j I have been modified as a part of a
 CC preformulation scheme to develop an optimised drug product for
 CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j I,
 CC modified and unmodified, are given in AAR82491-R82525. This peptide
 CC fragment corresponds to amino acids 71-90 of the allergen mature
 CC protein.

SO Sequence 20 AA;

Query Match 100.0%; Score 79; DB 15; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFSGNNNIKLKMPY 15
 DB 1 IFSGNNNIKLKMPY 15

RESULT 3

AA44682 AAM44682 standard; peptide; 30 AA.

AC AAM44682;

DT 01-MAY-1998 (first entry)

DE T-cell epitope peptide #1 of sugi pollen antigen.

KM T-cell epitope; sugi pollen antigen; sugi pollinosis.

XX T-cell epitope; sugi pollen antigen; sugi pollinosis.

OS Synthetic.

PN JP10007700-A.

PD 13-JAN-1998.

PF 24-JUN-1996; 96JP-0163287.

PR 24-JUN-1996; 96JP-0163287.

PA (DAIL) DAICEL CHEM IND LTD.

PA (MEIT) MEIJI SEIKA KAISHA LTD.

DR WPI; 1998-133630/13.

XX T cell epitope peptide of sugi pollen antigen - useful in the

PT treatment of sugi pollinosis

XX Claim 1; Page 1; 14pp; Japanese.

CC T-cell epitope peptides AAM44682-88 and their derivatives react with

CC sugi pollinosis patient peripheral blood T lymphocytes. A composition

CC prepared by combining at least 2 of the above peptides and/or their

CC derivatives is used for the prevention and treatment of sugi

CC pollinosis.

SO Sequence 30 AA;

OY 1 IFSGNNNIKLKMPY 15
 DB 11 IFSGNNNIKLKMPY 25

RESULT 4

AA27371 AAM27371 standard; peptide; 134 AA.

AC AAM27371;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #3.

KM Multi-epitope peptide; immunotherapeutic agent; allergic disease;

XX T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN WO9732600-A1.

PD 12-SEP-1997.
 XX PF 10-MAR-1997; 97WO-JP00740.
 XX PR 10-MAR-1996; 96JP-0080702.
 XX PA (MEIP) MEIJI MILK PROD CO LTD.
 XX PI Dairiki K, Iwama A, Kino K, Kume A, Some T;
 XX DR WPI; 1997-470495/43.
 XX PT Peptide immuno-therapeutic agent to treat allergic diseases -
 XX PT contains multi-epitope peptide containing T cell epitope regions
 XX PT from different allergens
 XX PS Claim 6; Page 32; 58pp; Japanese.
 XX CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX SQ Sequence 134 AA;
 SQ Query Match 100.0%; Score 79; DB 18; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IFSGNNNIIKLKMPY 15
 DB 78 IFSGNNNIIKLKMPY 92
 RESULT 5
 AAR75388
 ID AAR75388 standard; protein; 353 AA.
 XX AAR75388;
 XX 12-MAR-1996 (first entry)
 XX DE Japanese cedar pollen allergen Cryj I.
 XX KW Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides;
 XX KM prevention; treatment; cryptomeria pollinosis.
 XX OS Cryptomeria japonica.
 XX FH Key * Location/Qualifiers
 FT Peptide 61..75 /note= "T-cell epitope peptide"
 FT Peptide 91..105 /note= "T-cell epitope peptide"
 FT Peptide 106..120 /note= "T-cell epitope peptide"
 FT Peptide 146..160 /note= "T-cell epitope peptide"
 FT Peptide 211..225 /note= "T-cell epitope peptide"
 FT Peptide 326..340 /note= "T-cell epitope peptide"
 FT Peptide 335..346 /note= "T-cell epitope peptide"
 FT Peptide /note= "T-cell epitope peptide"
 XX JP07118295-A.

PD 09-MAY-1995.
 XX PF 20-OCT-1993; 93JP-0262626.
 XX PR 20-OCT-1993; 93JP-0262626.
 XX PA (MEIP) MEIJI MILK PROD CO LTD.
 XX DR WPI; 1995-203834/27.
 XX PT New cryptomeria pollen allergen T-cell epitope peptide - used for
 XX PT prevention, treatment and investigation of Japanese cedar pollinosis
 XX PS Disclosure; Figs 1-2; 8pp; Japanese.
 XX CC AAR75388 is the Japanese cedar pollen allergen Cryj I, from which the
 CC T-cell epitope peptides AAR89289-R89295 were derived. The peptides
 CC can be used for the prevention and treatment of cryptomeria
 CC pollinosis, and also for the investigation of pollinosis.
 XX SQ Sequence 353 AA;
 SQ Query Match 100.0%; Score 79; DB 16; Length 353;
 Best Local Similarity 100.0%; Pred. No. 3.2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IFSGNNNIIKLKMPY 15
 DB 71 IFSGNNNIIKLKMPY 85
 RESULT 6
 AAR81587
 ID AAR81587 standard; protein; 353 AA.
 XX AAR81587;
 XX 24-MAY-1996 (first entry)
 XX DE Cedar pollen allergen B.
 XX KW Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
 XX KM antibody; pollinosis; therapy; immunotherapy.
 XX OS Cryptomeria japonica.
 XX FH EP700929-A2.
 XX PN 13-MAR-1996.
 XX FD 08-SEP-1995; 95EP-0306295.
 XX PF 14-JUL-1995; 95JP-0200221.
 XX PR 10-SEP-1994; 94JP-0242137.
 XX PR 14-JUL-1995; 95JP-0200204.
 XX PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX PI Hino K, Saito S, Taniguchi Y;
 XX DR WPI; 1996-140976/15.
 XX PT New peptide(s) derived from cedar pollen allergens - activate
 XX PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
 XX PT used for treating cedar pollinosis
 XX PS Claim 5; Page 31-32; 36pp; English.
 XX CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (AAR81586) and B (AAR81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (AAR81573-79) essential

CC for T-cell recognition, and homologous peptides (AAR81588-96) can
CC be used as immunotherapeutic agents to treat or prevent cedar
CC pollinosis, avoiding side-effects such as anaphylaxis.

SO Sequence 353 AA;

Query Match 100.0%; Score 79; DB 17; Length 353;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFSGNNIKLKMPY 15
DB 71 IFSGNNIKLKMPY 85

RESULT 7
ID AAY25664
AAY25664 standard; protein; 373 AA.

AC AAY25664;

DT 30-SEP-1999 (first entry)

DE Cedar allergen 493634 Cry j IB protein fragment.

OS Major histocompatibility complex; class II; desensitizing; human;
KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

XX Cedrus sp.

XX WO93934826-A1.

XX 15-JUL-1999.

XX 11-JAN-1999; 99WO-GB00080.

XX 21-SEP-1998; 98GB-0020474.

XX 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Kay AB, Larche M;

XX WPI, 1999-458255/38.

XX Desensitizing patients to polypeptide allergens

XX Example 6, Page 73; 117pp; English.

CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenebrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a cedar (Cedrus sp.) allergen 493634 Cry j IB.

XX Sequence 373 AA;

Query Match. 100.0%; Score 79; DB 20; Length 373;

Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFSGNNIKLKMPY 15
DB 91 IFSGNNIKLKMPY 105

RESULT 8
ID AAY25668
AAY25668 standard; protein; 373 AA.

AC AAY25668;

DT 30-SEP-1999 (first entry)

DE Japanese cedar allergen 541803 Cry j I precursor protein fragment.

OS Major histocompatibility complex; class II; desensitizing; human;
KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

XX Cedrus sp.

XX WO93934826-A1.

XX 15-JUL-1999.

XX 11-JAN-1999; 99WO-GB00080.

XX 21-SEP-1998; 98GB-0020474.

XX 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Kay AB, Larche M;

XX WPI, 1999-458255/38.

XX Desensitizing patients to polypeptide allergens

XX Example 6, Page 75; 117pp; English.

CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenebrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a Japanese cedar (Cedrus sp.) allergen 541803 Cry j I
CC precursor.

XX Sequence 373 AA;

Query Match 100.0%; Score 79; DB 20; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFSGNNIKLKMPY 15
DB 91 IFSGNNIKLKMPY 105


```

RESULT 9
AAR31937
ID AAR31937 standard; Protein; 374 AA.
XX
AC AAR31937;
XX
DT 03-JUN-1993 (first entry)
XX
DE Cry j I.
XX
KM Japanese cedar pollen; allergen; antigen; allergy; B cell; T cell.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "signal peptide"
FT Protein 22..374
FT /note= "mature Cry j I"
XX
PN WO9301213-A.
XX
PD 21-JAN-1993.
XX
PF 10-JUL-1992; 92WO-US05661.
XX
PR 12-JUL-1991; 91US-0729134.
XX
PR 15-JUL-1991; 91US-0730452.
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PI Bond JF, Griffith J, Pollock J;
XX
DR WPI: 1993-045434/05.
XX
DR N-PSDB; AAQ5304.
XX
PT Nucleic acid sequence encoding Cryptomeria japonica allergen -
PT for the diagnosis treatment and prevention of allergic reactions
PT to Japanese cedar pollen
XX
PS Claim 11; Page 42; 69pp; English.
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 79; DB 14; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
    |||||
DB 92 IFSGNNNIKLKMPY 106

RESULT 10
AAR45541
ID AAR45541 standard; Protein; 374 AA.
XX
AC AAR45541;
XX
DT 13-JUL-1994 (first entry)
XX
DE Cry j I pollen allergen.

```

```

XX
KM Japanese cedar; detection; allergy; treatment; diagnosis;
XX
KM T cell epitope; sensitivity.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "signal peptide"
FT Peptide 22..374
FT /note= "mature peptide"
XX
PN WO9401560-A.
XX
PD 20-JAN-1994.
XX
PF 15-JAN-1993; 93WO-US00139.
XX
PR 01-SEP-1992; 92US-0938990.
XX
PR 10-JUL-1992; 92WO-US05661.
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PI Bond JF, Garman RD, Griffith J, Kuo M, Pollock J;
XX
DR WPI: 1994-035066/04.
XX
DR N-PSDB; AAQ55271.
XX
PT Antigens derived from Japanese cedar pollen allergen Cry j I -
PT contain at least two T cell epitope(s), used to treat or diagnose
PT allergy
XX
PS Disclosure; Fig 4; 137pp; English.
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 79; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
    |||||
DB 92 IFSGNNNIKLKMPY 106

RESULT 11
AAR60166
ID AAR60166 standard; Protein; 374 AA.
XX
AC AAR60166;
XX
DT 24-MAR-1995 (first entry)
XX
DE Japanese cedar pollen antigen CryjI.
XX
DE Japanese cedar; pollen antigen; allergen; CryjI; sugi; pollinosis.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /label= signal_peptide
FT Protein 22..374
FT /label= mature_CryjI
XX
PN JP06197768-A.

```

```

XX 19-JUL-1994.
PD
XX
XX 07-JAN-1993; 93JP-0001116.
PF
XX 07-JAN-1993; 93JP-0001116.
PR
XX (MEIJ) MEIJI SEIKA KAISHA.
PA
XX WPI; 1994-268680/33.
DR N-PSDB; AAQ71601.
DR
XX
XX Sugi (Japanese cedar) pollen antigen CryjI - is useful for
PT diagnosis, treatment and prevention of sugi pollinosis
PT
XX Claim 2; Page 5-7; 9pp; Japanese.
PS
XX The coding sequence for the Japanese cedar ("sugi") pollen allergen
CC CryjI was isolated from a cDNA library prepared from polyA mRNA. All
CC or part of the CryjI protein can be used for diagnosis, treatment
CC and prevention of sugi pollinosis.
CC
SQ Sequence 374 AA;
Query Match 100.0%; Score 79; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 IFSGNNNIIKLKMPY 15
| | | | | | | | | | | | | | | |
Db 92 IFSGNNNIIKLKMPY 106
| | | | | | | | | | | | | | | |
RESULT 12
AAR82490
ID AAR82490 standard; Protein; 374 AA.
XX
XX AAR82490;
AC
XX 15-APR-1996 (first entry)
DT
XX
XX Cry j I Japanese Cedar pollen allergen.
DE
XX Cry j I; Japanese cedar pollen allergen; modified; drug production;
KM allergy; Crytpomeria japonica.
KM
XX Crytpomeria japonica.
OS
XX WO9527786-A1.
PN
XX 19-OCT-1995.
PD
XX 06-APR-1995; 95WO-US04249.
PF
XX 06-DEC-1994; 94US-0350225.
PR 08-APR-1994; 94US-0226248.
PR
XX (IMMU-) IMMUNOLOGIC PHARM CORP.
PA
XX Chen X, Evans S, Franzen HM, Kuo M, Powers SP;
PI Shaked Z;
PI
XX WPI; 1995-166391/47.
DR N-PSDB; AAT04248.
DR
XX Modified Crytpomeria japonica (Cry j) I peptide(s) - useful for
PT treating allergy to Japanese cedar pollen allergen or
PT immunologically cross reactive allergens
PT
XX Disclosure; Figure 1, 60pp; English.
PS
XX Novel peptides of cry j I have been modified as a part of a
CC preformulation scheme to develop an optimised drug product for
CC

```

```

CC therapeutic treatment of humans suffering from allergy to Japanese
CC cedar pollen allergen or an allergen which is immunologically cross
CC reactive with Japanese cedar pollen allergen. Such modified peptides
CC possess certain characteristics which render them particularly
CC suitable for drug product formulation. Peptide fragments of Cry j I,
CC modified and unmodified, are given in AAR82491-R82525.
CC
SQ Sequence 374 AA;
Query Match 100.0%; Score 79; DB 16; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 IFSGNNNIIKLKMPY 15
| | | | | | | | | | | | | | | |
Db 92 IFSGNNNIIKLKMPY 106
| | | | | | | | | | | | | | | |
RESULT 13
AAY25665
ID AAY25665 standard; protein; 374 AA.
XX
XX AAY25665;
AC
XX 30-SEP-1999 (first entry)
DT
XX
XX Cedar allergen 493632 Cry j IA protein fragment.
DE
XX
XX Major histocompatibility complex; class II; desensitising; human;
KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX
XX Cedrus sp.
OS
XX WO9934826-A1.
PN
XX 15-JUL-1999.
PD
XX 11-JAN-1999; 99WO-GE00080.
PF
XX 21-SEP-1998; 98GB-0020474.
PR 09-JAN-1998; 98GB-0000445.
PR
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
PA
XX Kay AB, Larche M;
PI
XX WPI; 1999-458255/38.
DR
XX Desensitizing patients to polypeptide allergens
DR
XX Example 6; Page 73; 117pp; English.
PS
XX This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitising patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenbrilio mollitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a cedar (Cedrus sp.) allergen 493632 Cry j IA.
XX

```

SQ Sequence 374 AA;
 Query Match 100.0%; Score 79; DB 20; Length 374;
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
 |||||
 DB 92 IFSGNNNIKLKMPY 106

RESULT 14
 AAY255669
 ID AAY255669 standard; protein; 374 AA.
 XX
 AC AAY255669;
 XX
 DT 30-SEP-1999 (first entry)
 XX
 DE Japanese cedar allergen 541802 Cry j I precursor protein fragment.
 XX
 KW Major histocompatibility complex; class II; desensitizing; human;
 KW allergen; grass; tree; weed; pollen; fungi; mold; food; insect; sting;
 KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX
 OS Cedrus sp.
 XX
 PN WO934826-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 11-JAN-1999; 99WO-GB00080.
 XX
 PR 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Kay AB, Larche M;
 XX
 DR WPI; 1999-458255/38.
 XX
 PT Desensitizing patients to polypeptide allergens
 PS
 XX
 XX

Example 6; Page 75; 117pp; English.

This invention describes a novel method of desensitizing a patient to a polypeptide allergen and comprises administering to the patient a peptide derived from the allergen where restriction to a MHC Class II molecule possessed by the patient can be demonstrated for the peptide and the peptide is able to induce a late phase response in an individual who possesses the MHC Class II molecule. The methods can be used for desensitizing patients to allergens present in e.g. grass, tree and weed (including ragweed) pollens, fungi and moulds, foods, stinging insects, the chironomidae (non-biting midges), spiders and mites, housefly, fruit fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee, non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to produce immunological vaccines which may be used to prevent and/or treat conditions involving hypersensitivity to allergens. This sequence represents a Japanese cedar (Cedrus sp.) allergen 541802 Cry j I precursor.

SQ Sequence 374 AA;
 Query Match 100.0%; Score 79; DB 20; Length 374;
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
 |||||
 DB 92 IFSGNNNIKLKMPY 106

RESULT 15
 AAR45577
 ID AAR45577 standard; Protein; 367 AA.
 XX
 AC AAR45577;
 XX
 DT 13-JUL-1994 (first entry)
 XX
 DE Jun s I.
 XX
 KW Japanese cedar; pollen allergen; allergy; treatment; diagnosis;
 KW T cell epitope; sensitivity; detection.
 XX
 OS Juniperus sabinoides.

Key Location/Qualifiers
 FH Peptide 1..21 /note= "signal peptide"
 FT Peptide 22..367 /note= "mature peptide"

WO9401560-A.
 XX
 PD 20-JAN-1994.
 XX
 PF 15-JAN-1993; 93WO-US00139.
 XX
 PR 01-SEP-1992; 92US-0938990.
 PR 10-JUL-1992; 92WO-US05661.
 XX
 PA (IMMU-) IMMUTOLOGIC PHARM CORP.
 XX
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J;
 XX
 DR WPI; 1994-035066/04.
 DR N-P8DB; AAQ55272.
 XX
 XX

PT Antigens derived from Japanese cedar pollen allergen Cry j I -
 PT contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy
 XX

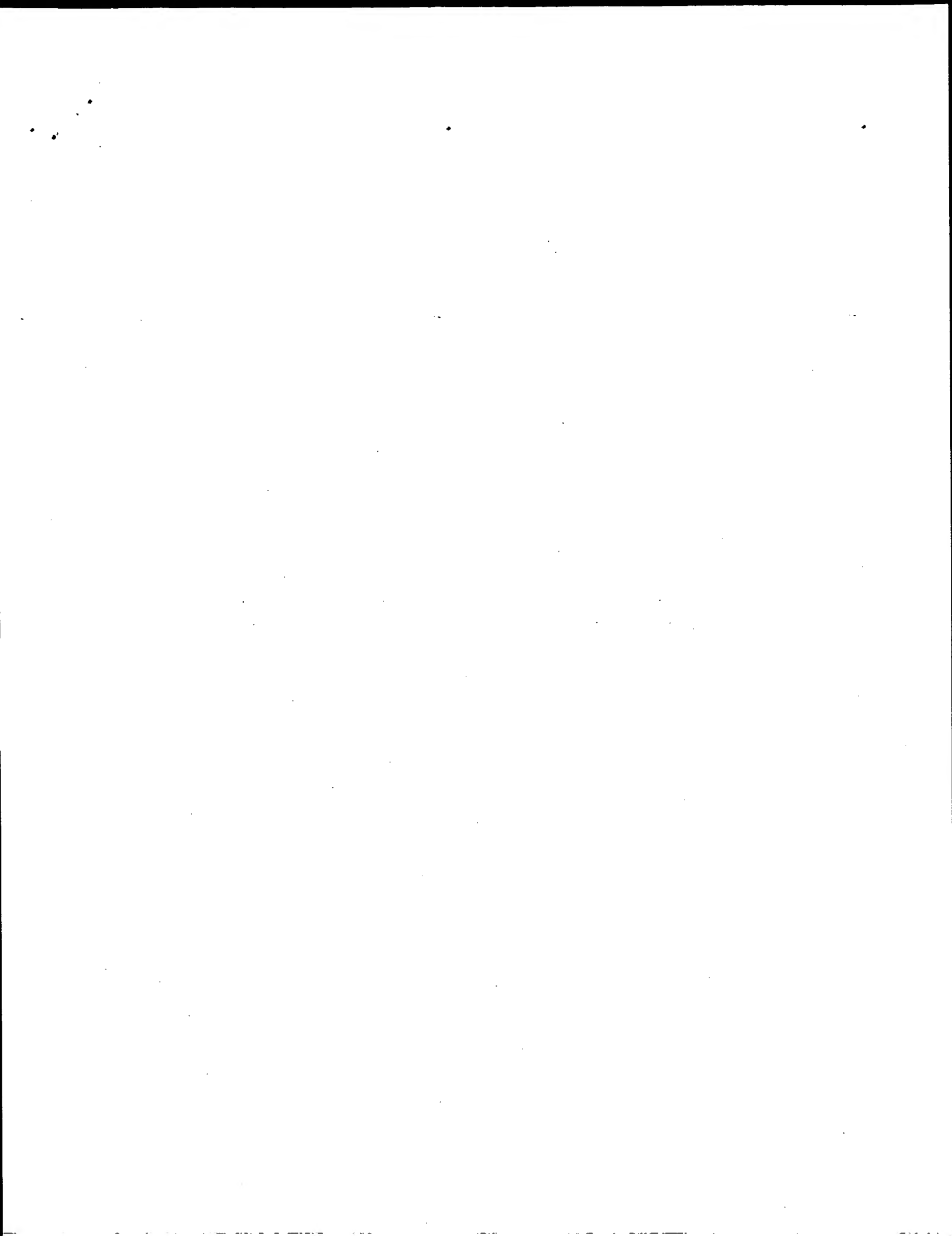
PS Disclosure; Fig 16; 137pp; English.

The sequence is that of Jun s I, a homologue of the Japanese cedar pollen allergen Cry j I. Antigenic peptides derived from it can be used for the treatment and diagnosis of allergies associated with Japanese cedar pollen.

SQ Sequence 367 AA;
 Query Match 86.1%; Score 68; DB 15; Length 367;
 Best Local Similarity 86.7%; Pred. No. 5.2e-05;
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IFSGNNNIKLKMPY 15
 |||||
 DB 92 IFSGNNNIKLKMPY 106

Search completed: April 20, 2003, 13:06:06
 Job time : 20.1974 secs



GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds
(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524D-30

Perfect score: 79

Sequence: 1 MNITLKMPTIAGYK 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: A_Geneseq 101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	20	AA45549	Cry j I pollen all
2	79	100.0	20	AA82498	Cry j I Japanese C
3	79	100.0	30	AA44682	T-cell epitope pep
4	79	100.0	134	AA27372	Multi-epitope pep
5	79	100.0	353	AA75386	Japanese cedar pol
6	79	100.0	353	AA81587	Japanese cedar pol
7	79	100.0	373	AA25664	Cedar pollen aller
8	79	100.0	373	AA25668	Japanese cedar all
9	79	100.0	374	AA31937	Cry j I. Cryptome
10	79	100.0	374	AA45541	Cry j I pollen all

11	79	100.0	374	AA60166	Japanese cedar pol
12	79	100.0	374	AA82490	Cry j I Japanese C
13	79	100.0	374	AA25665	Cedar allergen 493
14	79	100.0	374	AA25669	Japanese cedar all
15	70	88.6	357	AA45577	Japanese cedar all
16	70	88.6	370	AA45578	Japanese cedar all
17	63	79.7	105	AA27370	Multi-epitope pep
18	59	74.7	105	AA27370	T-cell epitope pep
19	59	74.7	31	AA27372	Multi-epitope pep
20	59	74.7	354	AA04344	Chamaecyparis obtu
21	59	74.7	354	AA42121	Japanese cypress p
22	59	74.7	375	AA04345	Chamaecyparis obtu
23	55	69.6	20	AA45550	Cry j I pollen all
24	55	69.6	20	AA82499	Cry j I Japanese C
25	55	69.6	30	AA45579	Cry j I pollen all
26	45	57.0	24	AA45580	Cry j I pollen all
27	41	51.9	607	AA23338	Drosophila Na+ dri
28	41	51.9	773	AA23336	Drosophila Na+ dri
29	41	51.9	774	AA23340	Drosophila Na+ dri
30	41	51.9	779	AA23341	Drosophila Na+ dri
31	41	51.9	779	AA23342	Drosophila Na+ dri
32	41	51.9	1017	AA860785	Drosophila melanog
33	41	51.9	1030	AA23337	Drosophila Na+ dri
34	41	51.9	1030	AA23339	Drosophila Na+ dri
35	41	51.9	1086	AA860786	Drosophila melanog
36	39	49.4	97	ABP01586	Human ORFX protein
37	39	49.4	162	AA38669	Neisseria meningit
38	38	48.1	20	AA42131	T-cell epitope pep
39	38	48.1	42	AA82824	Human immune haema
40	38	48.1	320	AA85420	Lactococcus lactis
41	38	48.1	444	AA22	Drosophila melanog
42	38	48.1	551	AA52398	Human keratin KRT
43	38	48.1	551	AA20423	Human keratin-2 (k
44	37	46.8	74	AA45475	Human secreted pro
45	37	46.8	104	AA004575	Human polypeptide

ALIGNMENTS

RESULT 1	AA45549	AA45549 standard; Protein; 20 AA.
ID	AA45549	
XX	AA45549	
AC	AA45549	
XX		
DT	13-JUL-1994	(first entry)
XX		
DE	Cry j I pollen allergen peptide Cyt-8.	
XX		
KW	Japanese cedar; detection; allergy; treatment; diagnosis;	
XX	T cell epitope; sensitivity.	
OS	Cryptomeria-japonica.	
XX		
PY	WO9401560-A.	
XX		
RD	20-JAN-1994.	
XX		
PF	13-JAN-1993	93WO-US00139.
XX		
PR	01-SEP-1992	92US-0938990.
XX		
PR	10-JUL-1992	92WO-US05661.
XX		
PA	(IMMU-) IMMUNOLOGIC PHARM CORP.	
XX		
PI	Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;	
XX		
DR	WPI; 1994-035066/04.	
XX		
PT	Antigens derived from Japanese cedar pollen allergen Cry j I -	
PT	contain at least two T cell epitope(s), used to treat or diagnose	
PT	allergy	

XX Claim 1; Fig 13; 137bp; English.
 PS
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j I (amino acids 71-90). The peptide, CJI-8,
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen. It has enhanced therapeutic properties
 CC but reduced side effects compared to naturally occurring allergens.
 CC
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 79; DB 15; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.3e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MNIKLKMPYIAGYK 15
 6 MNIKLKMPYIAGYK 20
 DB
 RESULT 2
 AAR82498 standard; Protein; 20 AA.
 ID AAR82498 standard; Protein; 20 AA.
 XX
 AC AAR82498;
 XX
 DT 15-APR-1996 (first entry)
 XX
 DE Cry j I Japanese Cedar pollen allergen peptide fragment (CJI-8).
 XX
 KW Cry j I; Japanese cedar pollen allergen; modified; drug production;
 XX allergy; Cryptomeria japonica.
 OS
 XX Cryptomeria japonica.
 PN WO9527786-A1.
 XX
 PD 19-OCT-1995.
 XX
 PF 06-APR-1995; 95WO-US04249.
 XX
 PR 06-DEC-1994; 94US-0350225.
 PR 08-APR-1994; 94US-0226248.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP;
 PI Shaked Z;
 XX
 DR WPI, 1995-366391/47.
 XX
 PT Modified Cryptomeria japonica (Cry j) I peptide(s) - useful for
 PT treating allergy to Japanese cedar pollen allergen or
 PT immunologically cross reactive allergens
 XX
 PS Disclosure; Figure 2; 60bp; English.
 XX
 CC Novel peptides of cry j I have been modified as a part of a
 CC preformulation scheme to develop an optimised drug product for
 CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j I,
 CC modified and unmodified, are given in AAR82491-R82525. This peptide
 CC fragment corresponds to amino acids 71-90 of the allergen mature
 CC protein.
 CC
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 79; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.3e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNIKLKMPYIAGYK 15
 DB 6 MNIKLKMPYIAGYK 20
 DB
 RESULT 3
 AAW44682 standard; peptide; 30 AA.
 ID AAW44682 standard; peptide; 30 AA.
 XX
 AC AAW44682;
 XX
 DT 01-MAY-1998 (first entry)
 XX
 DE T-cell epitope peptide #1 of sugi pollen antigen.
 XX
 KW T-cell epitope; sugi pollen antigen; sugi pollinosis.
 XX
 OS Synthetic.
 OS Cryptomeria japonica.
 XX
 PN JP10007700-A.
 XX
 PD 13-JAN-1998.
 XX
 PE 24-JUN-1996; 96UP-0163287.
 XX
 FR 24-JUN-1996; 96UP-0163287.
 XX
 PA (DAI) DAICEL CHEM IND LTD.
 PA (MEIJ) MEIJI SEIKA KAISHA LTD.
 XX
 DR WPI, 1998-133630/13.
 XX
 PT T cell epitope peptide of sugi pollen antigen - useful in the
 PT treatment of sugi pollinosis
 XX
 PS Claim 1; Page 1; 14pp; Japanese.
 XX
 CC T-cell epitope peptides AAW44682-88 and their derivatives react with
 CC sugi pollinosis patient peripheral blood T lymphocytes. A composition
 CC prepared by combining at least 2 of the above peptides and/or their
 CC derivatives is used for the prevention and treatment of sugi
 CC pollinosis.
 CC
 SQ Sequence 30 AA;
 Query Match 100.0%; Score 79; DB 19; Length 30;
 Best Local Similarity 100.0%; Pred. No. 6.9e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MNIKLKMPYIAGYK 15
 DB 16 MNIKLKMPYIAGYK 30
 DB
 RESULT 4
 AAW27371 standard; peptide; 134 AA.
 ID AAW27371 standard; peptide; 134 AA.
 XX
 AC AAW27371;
 XX
 DT 24-MAR-1998 (first entry)
 XX
 DE Multi-epitope peptide used as immunotherapeutic agent #3.
 XX
 DE Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 XX
 OS Synthetic.
 OS WO9732600-A1.
 XX
 PN

PD 12-SEP-1997.
 XX
 PF 10-MAR-1997; 97WO-JP00740.
 XX
 PR 10-MAR-1996; 96JP-0080702.
 XX
 PA (MEIP) MEIJI MILK PROD CO LTD.
 XX
 PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;
 XX
 DR WPI; 1997-470495/43.
 PT Peptide immuno-therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens
 XX
 PS Claim 6; Page 32; 58pp; Japanese.
 XX
 CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen, have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen, and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX
 SQ Sequence 134 AA;

Query Match 100.0%; Score 79; DB 18; Length 134;
 Best Local Similarity 100.0%; Pred. No. 3.9e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNIKLKMPTYIAGYK 15
 DB 83 MNIKLKMPTYIAGYK 97

RESULT 5
 AAR75388
 ID AAR75388 standard; protein, 353 AA.
 XX
 AC AAR75388;
 XX
 DT 12-MAR-1996 (first entry)
 XX
 DE Japanese cedar pollen allergen Cryj I.
 XX
 KM Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides;
 KM prevention; treatment; cryptomeria pollinosis.
 XX
 OS Cryptomeria japonica.
 XX
 FH Key
 FH Peptide Location/Qualifiers
 FT /note= "T-cell epitope peptide"
 FT 61..75
 FT Peptide 91..105
 FT /note= "T-cell epitope peptide"
 FT 106..120
 FT Peptide /note= "T-cell epitope peptide"
 FT 146..160
 FT Peptide /note= "T-cell epitope peptide"
 FT 211..225
 FT Peptide /note= "T-cell epitope peptide"
 FT 326..340
 FT Peptide /note= "T-cell epitope peptide"
 FT 335..346
 FT Peptide /note= "T-cell epitope peptide"
 XX
 PN JP07118295-A.

PD 09-MAY-1995.
 XX
 PF 20-OCT-1993; 93JP-0262626.
 XX
 PR 20-OCT-1993; 93JP-0262626.
 XX
 PA (MEIP) MEIJI MILK PROD CO LTD.
 XX
 PI WPI; 1995-203834/27.
 XX
 DR New cryptomeria pollen allergen T-cell epitope peptide - used for
 PT prevention, treatment and investigation of Japanese cedar pollinosis
 PT Disclosure; Figs 1-2; 8pp; Japanese.
 XX
 PS AAR75388 is the Japanese cedar pollen allergen Cryj I, from which the
 CC T-cell epitope peptides AAR89289-R89295 were derived. The peptides
 CC can be used for the prevention and treatment of cryptomeria
 CC pollinosis, and also for the investigation of pollinosis.
 XX
 SQ Sequence 353 AA;

Query Match 100.0%; Score 79; DB 16; Length 353;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNIKLKMPTYIAGYK 15
 DB 76 MNIKLKMPTYIAGYK 90

RESULT 6
 AAR81587
 ID AAR81587 standard; protein, 353 AA.
 XX
 AC AAR81587;
 XX
 DT 24-MAY-1996 (first entry)
 XX
 DE Cedar pollen allergen B.
 XX
 KM Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
 KM antibody; pollinosis; therapy; immunotherapy.
 XX
 OS Cryptomeria japonica.
 XX
 EN EP700929-A2.
 XX
 PD 13-MAR-1996.
 XX
 PF 08-SEP-1995; 95EP-0306295.
 XX
 PR 14-JUL-1995; 95JP-0200221.
 PR 10-SEP-1994; 94JP-0242137.
 PR 14-JUL-1995; 95JP-0200204.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 PI Hino K, Saito S, Taniguchi Y;
 XX
 DR WPI; 1996-140976/15.
 XX
 PT New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
 PT used for treating cedar pollinosis
 XX
 PS Claim 5; Page 31-32; 36pp; English.
 XX
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (AAR81586) and B (AAR81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (AAR81573-79) essential

CC for T-cell recognition, and homologous peptides (AAR81588-96) can
CC be used as immunotherapeutic agents to treat or prevent cedar
CC pollinosis, avoiding side-effects such as anaphylaxis.

SQ Sequence 353 AA;

Query Match 100.0%; Score 79; DB 17; Length 353;

Best Local Similarity 100.0%; Pred. No. 1.2e-06; Mismatches 0; Gaps 0;

Matches 15; Conservative 0; Indels 0; Gaps 0;

OY 1 MNIKLKMPTIAGYK 15

Db 76 MNIKLKMPTIAGYK 90

RESULT 7
ID AAY25664 standard; protein; 373 AA.

AC AAY25664;

DT 30-SEP-1999 (first entry)

DE Cedar allergen 493634 Cry j IB protein fragment.

KM Major histocompatibility complex; class II; desensitizing; human;
KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.

PN WO9934826-A1.

PD 15-JUL-1999.

PF 11-JAN-1999; 99WO-GB00080.

PR 21-SEP-1998; 98GB-0020474.

PR 09-JAN-1998; 98GB-0000445.

PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

PI Kay AB, Larche M;

DR WPI; 1999-458255/38.

PT Desensitizing patients to polypeptide allergens

PS Example 6; Page 73; 117pp; English.

CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a cedar (Cedrus sp.) allergen 493634 Cry j IB.

SQ Sequence 373 AA;

Query Match 100.0%; Score 79; DB 20; Length 373;

Best Local Similarity 100.0%; Pred. No. 1.3e-06; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNIKLKMPTIAGYK 15

Db 96 MNIKLKMPTIAGYK 110

RESULT 8

ID AAY25668 standard; protein; 373 AA.

AC AAY25668;

DT 30-SEP-1999 (first entry)

DE Japanese cedar allergen 541803 Cry j I precursor protein fragment.

KM Major histocompatibility complex; class II; desensitizing; human;
KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.

PN WO9934826-A1.

PD 15-JUL-1999.

PF 11-JAN-1999; 99WO-GB00080.

PR 21-SEP-1998; 98GB-0020474.

PR 09-JAN-1998; 98GB-0000445.

PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

PI Kay AB, Larche M;

DR WPI; 1999-458255/38.

PT Desensitizing patients to polypeptide allergens

PS Example 6; Page 75; 117pp; English.

CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a Japanese cedar (Cedrus sp.) allergen 541803 Cry j I

SQ Sequence 373 AA;

Query Match 100.0%; Score 79; DB 20; Length 373;

Best Local Similarity 100.0%; Pred. No. 1.3e-06; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNIKLKMPTIAGYK 15

Db 96 MNIKLKMPTIAGYK 110


```

RESULT 9
AAR31937 standard; Protein; 374 AA.
ID AAR31937
XX
AC AAR31937;
XX
DT 03-JUN-1993 (first entry)
XX
DE Cry j I.
XX
KW Japanese cedar pollen; allergen; antigen; allergy; B cell; T cell.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "signal peptide"
FT Protein 22..374
FT /note= "mature Cry j I"
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PN WO9301213-A.
XX
PD 21-JAN-1993.
XX
PF 10-JUL-1992; 92WO-US05661.
XX
PR 12-JUL-1991; 91US-0729134.
PR 15-JUL-1991; 91US-0730452.
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PI Bond JF, Griffith J, Pollock J;
XX
DR WPI; 1993-045434/05.
DR N-PSDB; AAQ35304.
XX
PT Nucleic acid sequence encoding Cryptomeria japonica allergen -
PT for the diagnosis treatment and prevention of allergic reactions
PT to Japanese cedar pollen
XX
PS Claim 11; Page 42; 69pp; English.
XX
CC Fresh pollen and staminate cone samples were collected from a single
CC Cryptomeria japonica (Japanese cedar) tree. RNA was prep'd. and used
CC to synthesize cDNA. The cDNA was subjected to successive rounds of
CC PCR to yield a full length Cry j I clone. Cry j I or an antigenic
CC fragment of it may be used for detecting, treating and preventing an
CC allergic response to Japanese cedar pollen allergen. It is capable of
CC modifying both the B and T cell response to Cry j I and T cell response
CC to a Cry j I antigen.
XX
SQ Sequence 374 AA;
XX
Query Match 100.0%; Score 79; DB 14; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNIKLKMPTIAGYK 15
DB 97 MNIKLKMPTIAGYK 111
XX
RESULT 10
AAR45541
ID AAR45541 standard; Protein; 374 AA.
XX
AC AAR45541;
XX
DT 13-JUL-1994 (first entry)
XX
DE Cry j I pollen allergen.

```

```

XX
KW Japanese cedar; detection; allergy; treatment; diagnosis;
XX T cell epitope; sensitivity.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "signal peptide"
FT Peptide 22..374
FT /note= "mature peptide"
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PN WO9401560-A.
XX
PD 20-JAN-1994.
XX
PF 15-JAN-1993; 93WO-US00139.
XX
PR 01-SEP-1992; 92US-0938990.
PR 10-JUL-1992; 92WO-US05661.
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PI Bond JF, Garman RD, Griffith J, Kuo M, Pollock J;
XX
DR WPI; 1994-035066/04.
DR N-PSDB; AAQ55271.
XX
PT Antigens derived from Japanese cedar pollen allergen Cry j I -
PT contain at least two T cell epitope(s), used to treat or diagnose
PT allergy
XX
PS Disclosure; Fig 4; 137pp; English.
XX
CC The sequence is that of the Japanese cedar pollen allergen
CC Cry j I which contains at least two T cell epitopes. Peptide
CC antigens derived from it can be used for the treatment and
CC diagnosis of allergies associated with Japanese cedar pollen.
CC The peptides have enhanced therapeutic properties but reduced
CC side effects compared to naturally occurring allergens.
XX
SQ Sequence 374 AA;
XX
Query Match 100.0%; Score 79; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNIKLKMPTIAGYK 15
DB 97 MNIKLKMPTIAGYK 111
XX
RESULT 11
AAR60166
ID AAR60166 standard; Protein; 374 AA.
XX
AC AAR60166;
XX
DT 24-MAR-1995 (first entry)
XX
DE Japanese cedar pollen antigen CryjI.
XX
KW Japanese cedar; pollen antigen; allergen; CryjI; sugi; pollinosis.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /label= "signal_peptide"
FT Protein 22..374
FT /label= "mature_CryjI"
XX
PN JP06197768-A.

```

XX 19-JUL-1994.
 PD 07-JAN-1993; 93JP-0001116.
 XX 07-JAN-1993; 93JP-0001116.
 PR (MEIJ) MEIJI SEIKA KAISHA.
 XX N-PSDB; AAT04248.
 DR N-PSDB; AAT04248.
 XX Sugi (Japanese cedar) pollen antigen CryjI - is useful for
 PT diagnosis, treatment and prevention of sugi pollinosis
 XX Claim 2; Page 5-7; 9pp; Japanese.
 XX The coding sequence for the Japanese cedar ("sugi") pollen allergen
 CC CryjI was isolated from a cDNA library prepared from polyA mRNA. All
 CC or part of the CryjI protein can be used for diagnosis, treatment
 CC and prevention of sugi pollinosis.
 XX Sequence 374 AA;
 SQ
 Query Match 100.0%; Score 79; DB 15; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MNIKMKMPYIAGYK 15
 Db 97 MNIKMKMPYIAGYK 111
 RESULT 12
 AAR82490
 ID AAR82490 standard; Protein; 374 AA.
 XX AAR82490;
 AC
 XX 15-APR-1996 (first entry)
 DT
 XX
 XX Cry j I Japanese cedar pollen allergen.
 DE
 XX Cry j I; Japanese cedar pollen allergen; modified; drug production;
 KM allergy; Crytpomeria japonica.
 KW
 XX Crytpomeria japonica.
 OS
 XX MO9527786-A1.
 PN
 XX 19-OCT-1995.
 PD
 XX 06-APR-1995; 95WO-US04249.
 PF
 XX 06-DEC-1994; 94US-0350225.
 PR 08-APR-1994; 94US-0226248.
 XX (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX Chen X, Evans S, Franzen HM, Kuo M, Powers SP;
 PI Shaked Z;
 DR WPI; 1995-366391/47.
 XX N-PSDB; AAT04248.
 XX Modified Crytpomeria japonica (Cry j) I peptide(s) - useful for
 PT treating allergy to Japanese cedar pollen allergen or
 PT immunologically cross reactive allergens
 XX Disclosure; Figure 1; 60pp; English.
 PS Novel peptides of cry j I have been modified as a part of a
 CC preformulation scheme to develop an optimised drug product for

CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j I,
 CC modified and unmodified, are given in AAR82491-R82525.
 XX
 SQ Sequence 374 AA;
 Query Match 100.0%; Score 79; DB 16; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MNIKMKMPYIAGYK 15
 Db 97 MNIKMKMPYIAGYK 111
 RESULT 13
 AAY25665
 ID AAY25665 standard; protein; 374 AA.
 XX AAY25665;
 AC
 XX 30-SEP-1999 (first entry)
 DT
 XX Cedar allergen 493632 Cry j IA protein fragment.
 DE
 XX Major histocompatibility complex; class II; desensitizing; human;
 KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX
 XX Cedrus sp.
 OS
 XX WO9934826-A1.
 PN
 XX 15-JUL-1999.
 PD
 XX 11-JAN-1999; 99WO-GB00080.
 PF
 XX 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX Kay AB, Larche M;
 PI WPI; 1999-458255/38.
 DR Desensitizing patients to polypeptide allergens
 XX
 XX Example 6; Page 73; 117pp; English.
 PS
 XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The method can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a cedar (Cedrus sp.) allergen 493632 Cry j IA.

SQ Sequence 374 AA;
 Query Match 100.0%; Score 79; DB 20; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNIKLKMPMYIAGYK 15
 DB 97 MNIKLKMPMYIAGYK 111
 RESULT 14
 ID AAY25669 standard; protein; 374 AA.
 AC AAY25669;
 XX
 DT 30-SEP-1999 (first entry)
 DE Japanese cedar allergen 541802 Cry j I precursor protein fragment.
 XX
 KM Major histocompatibility complex; class II; desensitizing; human;
 KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; stinging;
 KM chironomidae; spider; mice; housefly; fruit fly; sheep blow fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 OS Cedrus sp.
 XX
 PN WO9934826-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 11-JAN-1999; 99WO-GB00080.
 XX
 PR 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Kay AB, Larche M;
 XX
 DR WPI; 1999-458255/38.
 XX
 PT Desensitizing patients to polypeptide allergens
 PS Example 6; Page 75; 117pp; English.
 XX
 CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrilio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC check immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 541802 Cry j I
 CC precursor.
 CC
 SQ Sequence 374 AA;
 XX
 Query Match 100.0%; Score 79; DB 20; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNIKLKMPMYIAGYK 15
 DB 97 MNIKLKMPMYIAGYK 111
 RESULT 15
 ID AAR45577 standard; Protein; 367 AA.
 AC AAR45577;
 XX
 DT 13-JUL-1994 (first entry)
 DE Jun s I.
 XX
 DE Japanese cedar; pollen allergen; allergy; treatment; diagnosis;
 KM T cell epitope; sensitivity; detection.
 KM
 XX Juniperus sabinoides.
 OS
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /note="signal peptide"
 FT Peptide 22..367
 FT /note="mature peptide"
 PN WO9401560-A.
 XX
 PD 20-JAN-1994.
 XX
 PF 15-JAN-1993; 93WO-US00139.
 XX
 PR 01-SEP-1992; 92US-0938990.
 PR 10-JUL-1992; 92WO-US05661.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J;
 XX
 DR WPI; 1994-035066/04.
 DR N-PSDB; AAQ55272.
 XX
 CC Antigen derived from Japanese cedar pollen allergen Cry j I -
 CC contain at least two T cell epitope(s), used to treat or diagnose
 CC allergy
 PS Disclosure; Fig 16; 137pp; English.
 XX
 CC The sequence is that of Jun s I, a homologue of the Japanese
 CC cedar pollen allergen Cry j I. Antigenic peptides derived from it
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen.
 CC
 SQ Sequence 367 AA;
 XX
 Query Match 88.6%; Score 70; DB 15; Length 367;
 Best Local Similarity 80.0%; Pred. No. 6.4e-05;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNIKLKMPMYIAGYK 15
 DB 97 MNIKLKMPDYVAGHK 111
 Search completed: April 20, 2003, 13:06:06
 Job time : 18.1974 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds
(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524D-31
Perfect score: 83
Sequence: 1 KMPMYTAGYKTFDGR 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	100.0	20	15	AA845550
2	83	100.0	20	15	AA845550
3	83	100.0	20	15	AA845550
4	83	100.0	30	15	AA845579
5	83	100.0	105	18	AAW27370
6	83	100.0	134	18	AAW27371
7	83	100.0	353	16	AA875388
8	83	100.0	353	17	AA881587
9	83	100.0	373	20	AA255664
10	83	100.0	374	14	AA31337

11	83	100.0	374	15	AA845541
12	83	100.0	374	15	AA860166
13	83	100.0	374	15	AA882490
14	83	100.0	374	20	AA255665
15	83	100.0	374	20	AA255669
16	73	88.0	24	15	AA845580
17	68	81.9	367	15	AA845577
18	68	81.9	370	15	AA845578
19	61	73.5	22	15	AA845581
20	61	73.5	24	15	AA845582
21	60	72.3	20	19	AAW42131
22	60	72.3	354	17	AAW04344
23	60	72.3	354	19	AAW04344
24	60	72.3	375	17	AAW04345
25	55	66.3	20	15	AA845549
26	55	66.3	20	16	AA882498
27	55	66.3	30	19	AAW46882
28	46	55.4	269	21	AA845512
29	46	55.4	269	21	AA845512
30	46	55.4	392	21	AA845511
31	46	55.4	392	21	AA845511
32	46	55.4	392	23	AA893832
33	46	55.4	410	21	AA823090
34	46	55.4	410	21	AA846510
35	46	55.4	411	22	AA868274
36	45	54.2	3508	22	AB864461
37	44	53.0	85	23	ABP33013
38	42	50.6	277	22	AB805062
39	42	50.6	277	23	AB847785
40	42	50.6	338	15	AA865965
41	41	49.4	103	21	AA849038
42	41	49.4	147	21	AA809173
43	41	49.4	181	21	AA809172
44	41	49.4	194	21	AA854925
45	41	49.4	203	21	AA809171

ALIGNMENTS

RESULT 1
ID AAR45550 standard; Protein; 20 AA.
XX AAR45550;
XX
XX
DT 13-JUN-1994 (first entry)
XX
XX Cry j I pollen allergen peptide CUI-9.
XX
XX Japanese cedar; detection; allergy; treatment; diagnosis;
XX T cell epitope; sensitivity.
XX
XX Cryptomeria japonica.
XX
XX W09401560-A.
XX
XX 20-JAN-1994.
XX
XX
XX 15-JAN-1993; 93WO-US00139.
XX
XX 01-SEP-1992; 92US-0938990.
XX 10-JUN-1992; 92WO-US05661.
XX
XX (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
XX Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;
XX WPI; 1994-035066/04.
XX
XX Antigen derived from Japanese cedar pollen allergen Cry j I -
XX contain at least two T cell epitope(s), used to treat or diagnose
XX allergy

XX PS Claim 1; Fig 13; 137pp; English.
 XX CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j I (amino acids 81-100). The peptide, CJI-9,
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen. It has enhanced therapeutic properties
 CC but reduced side effects compared to naturally occurring allergens.
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 83; DB 15; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KMPMYIAGYKTFDGR 15
 DB 1 KMPMYIAGYKTFDGR 15
 RESULT 2
 AAR82499 ID AAR82499 standard; Protein; 20 AA.
 XX AC AAR82499;
 XX DT 15-APR-1996 (first entry)
 XX DE Cry j I Japanese Cedar pollen allergen peptide fragment (CJI-9).
 XX KM Cry j I; Japanese cedar pollen allergen; modified; drug production;
 XX KW allergy; Crypomeria japonica.
 XX OS Crypomeria japonica.
 XX PN WO9527786-A1.
 XX PD 19-OCT-1995.
 XX PF 06-APR-1995; 95WO-US04249.
 XX PR 06-DEC-1994; 94US-0350225.
 XX PR 08-APR-1994; 94US-0226248.
 XX PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP;
 XX PI Shaked Z;
 XX DR WPI; 1995-366391/47.
 XX PT Modified Crypomeria japonica (Cry j) I peptide(s) - useful for
 XX PT treating allergy to Japanese cedar pollen allergen or
 XX PT immunologically cross reactive allergens
 XX PS Disclosure; Figure 2; 60pp; English.
 XX CC Novel peptides of cry j I have been modified as a part of a
 CC preformulation scheme to develop an optimised drug product for
 CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j I,
 CC modified and unmodified, are given in AAR82491-R82525. This peptide
 CC fragment corresponds to amino acids 81-100 of the allergen mature
 CC protein.
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 83; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMPMYIAGYKTFDGR 15
 DB 1 KMPMYIAGYKTFDGR 15
 RESULT 3
 AAR45579 ID AAR45579 standard; Protein; 30 AA.
 XX AC AAR45579;
 XX DT 13-JUL-1994 (first entry)
 XX DE Cry j I pollen allergen peptide CJI-41.
 XX KM Japanese cedar; detection; allergy; treatment; diagnosis;
 XX KW T cell epitope; sensitivity.
 XX OS Crypomeria japonica.
 XX PN WO9401560-A.
 XX PD 20-JAN-1994.
 XX PF 15-JAN-1993; 93WO-US00139.
 XX PR 01-SEP-1992; 92US-0938890.
 XX PR 10-JUL-1992; 92WO-US05661.
 XX PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J;
 XX DR WPI; 1994-035066/04.
 XX PT Antigens derived from Japanese cedar pollen allergen Cry j I -
 XX PT contain at least two T cell epitope(s), used to treat or diagnose
 XX PT allergy
 XX PS Claim 76; Fig 18; 137pp; English.
 XX CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j I. The peptide, CJI-41, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 XX SQ Sequence 30 AA;
 Query Match 100.0%; Score 83; DB 15; Length 30;
 Best Local Similarity 100.0%; Pred. No. 7.4e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KMPMYIAGYKTFDGR 15
 DB 1 KMPMYIAGYKTFDGR 15
 RESULT 4
 AAM27370 ID AAM27370 standard; peptide; 105 AA.
 XX AC AAM27370;
 XX DT 24-MAR-1998 (first entry)
 XX DE Multi-epitope peptide used as immunotherapeutic agent #2.
 XX KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 XX KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 XX OS Synthetic.

[illegible]

PS Disclosure; Figs 1-2; 8pp; Japanese.

CC AAR5388 is the Japanese cedar pollen allergen Cry I, from which the
CC T-cell epitope peptides AAR8289-R89295 were derived. The peptides
CC can be used for the prevention and treatment of cryptomera
CC pollinosis, and also for the investigation of pollinosis.

XX Sequence 353 AA;

SO Query Match 100.0%; Score 83; DB 16; Length 353;

Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMPMYIAGYKTFDGR 15
DB 81 KMPMYIAGYKTFDGR 95

RESULT 7

AAR81587 standard; Protein; 353 AA.

XX AAR81587;

XX 24-MAY-1996 (first entry)

XX Cedar pollen allergen B.

XX Cedar; pollen; allergen; immunoglobulin E; IGE; T-cell epitope;

XX antibody; pollinosis; therapy; immunotherapy.

XX Cryptomera japonica.

XX EP700929-A2.

XX 13-MAR-1996.

XX 08-SEP-1995; 95EP-0306295.

XX 14-JUL-1995; 95JP-0200221.

XX 10-SEP-1994; 94JP-0242137.

XX 14-JUL-1995; 95JP-0200204.

XX (HAYB) HAYASHIBARA SEIBUTSU KIGAKU.

XX Hino K, Saio S, Taniguchi Y;

XX WPI; 1996-140976/15.

XX Claim 5; Page 31-32; 36pp; English.

XX Synthetic peptides based on portions of cedar pollen allergens A

CC (AAR81586) and B (AAR81587) were tested for their ability to activate

CC cedar allergen-specific T-cells, but not allergen-specific IGE

CC antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell

CC epitopes. These peptides, plus subsequences (AAR81573-79) essential

CC for T-cell recognition, and homologous peptides (AAR81588-96) can

CC be used as immunotherapeutic agents to treat or prevent cedar

CC pollinosis, avoiding side-effects such as anaphylaxis.

XX Sequence 353 AA;

SO Query Match 100.0%; Score 83; DB 17; Length 353;

Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMPMYIAGYKTFDGR 15
DB 81 KMPMYIAGYKTFDGR 95

RESULT 8

AAV25664 standard; protein; 373 AA.

XX AAV25664;

XX 30-SEP-1999 (first entry)

XX Cedar allergen 493634 Cry j IB protein fragment.

XX Major histocompatibility complex; class II; desensitizing; human;

XX allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;

XX chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;

XX screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;

XX cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;

XX mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

XX Cedrus sp.

XX WO9934826-A1.

XX 15-JUL-1999.

XX 11-JAN-1999; 99WO-GB00080.

XX 21-SEP-1998; 98GB-0020474.

XX 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Kay AB, Larche M;

XX WPI; 1999-458255/38.

XX Desensitizing patients to polypeptide allergens

XX Example 6; Page 73; 117pp; English.

XX This invention describes a novel method of desensitizing a patient to a

CC polypeptide allergen and comprises administering to the patient a peptide

CC derived from the allergen where restriction to a MHC Class II molecule

CC possessed by the patient can be demonstrated for the peptide and the

CC peptide is able to induce a late phase response in an individual who

CC possesses the MHC Class II molecule. The methods can be used for

CC desensitizing patients to allergens present in e.g. grass, tree and weed

CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,

CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit

CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee, of

CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of

CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,

CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to

CC produce immunological vaccines which may be used to prevent and/or treat

CC conditions involving hypersensitivity to allergens. This sequence

XX Sequence 373 AA;

SO Query Match 100.0%; Score 83; DB 20; Length 373;

Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMPMYIAGYKTFDGR 15
DB 101 KMPMYIAGYKTFDGR 115

RESULT 9

AAV25668 standard; protein; 373 AA.

XX AAV25668;


```

XX 20-JAN-1994.
PD 15-JAN-1993; 93WO-US00139.
XX
PF 01-SEP-1992; 92US-0938990.
XX
PR 10-JUL-1992; 92WO-US05661.
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J;
XX WPI; 1994-035066/04.
DR N-PSDB; AAQ55271.
XX
PT Antigens derived from Japanese cedar pollen allergen Cry j I -
PT contain at least two T cell epitope(s), used to treat or diagnose
PT allergy
XX
PS Disclosure; Fig 4; 137pp; English.
XX
CC The sequence is that of the Japanese cedar pollen allergen
CC Cry j I which contains at least two T cell epitopes. Peptide
CC antigens derived from it can be used for the treatment and
CC diagnosis of allergies associated with Japanese cedar pollen.
CC The peptides have enhanced therapeutic properties but reduced
CC side effects compared to naturally occurring allergens.
XX
SQ Sequence 374 AA;
XX
Query Match 100.0%; Score 83; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KMPWYIAGYKTFDGR 15
DB 102 KMPWYIAGYKTFDGR 116
XX
RESULT 12
AAR60166
ID AAR60166 standard; Protein; 374 AA.
XX
AC AAR60166;
XX
DT 24-MAR-1995 (first entry)
XX
DE Japanese cedar pollen antigen CryjI.
XX
KW Japanese cedar; pollen antigen; allergen; CryjI; sugi; pollinosis.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /label= signal_peptide
FT Protein 22..374
FT /label= mature_CryjI
XX
XX JP06197768-A.
XX
XX 19-JUL-1994.
XX
XX 07-JAN-1993; 93JP-0001116.
XX
XX 07-JAN-1993; 93JP-0001116.
XX
XX (MEIJ ) MEIJI SEIKA KAISHA.
XX
XX WPI; 1994-268680/33.
XX
XX N-PSDB; AAQ71601.
XX
PT Sugi (Japanese cedar) pollen antigen CryjI - is useful for

```

```

PT diagnosis, treatment and prevention of sugi pollinosis
XX
XX Claim 2; Page 5-7; 9pp; Japanese.
XX
XX The coding sequence for the Japanese cedar ("sugi") pollen allergen
XX CryjI was isolated from a cDNA library prepared from polyA mRNA. All
XX or part of the CryjI protein can be used for diagnosis, treatment
XX and prevention of sugi pollinosis.
XX
SQ Sequence 374 AA;
XX
Query Match 100.0%; Score 83; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KMPWYIAGYKTFDGR 15
DB 102 KMPWYIAGYKTFDGR 116
XX
RESULT 13
AAR82490
ID AAR82490 standard; Protein; 374 AA.
XX
AC AAR82490;
XX
DT 15-APR-1996 (first entry)
XX
DE Cry j I Japanese Cedar pollen allergen.
XX
KW Cry j I; Japanese cedar pollen allergen; modified; drug production;
KW allergy; Cryptomeria japonica.
XX
OS Cryptomeria japonica.
XX
PN WO9527786-A1.
XX
PD 19-OCT-1995.
XX
PF 06-APR-1995; 95WO-US04249.
XX
PR 06-DEC-1994; 94US-0350225.
XX
PR 08-APR-1994; 94US-0226248.
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP,
PI Shaked Z;
XX
DR WPI; 1995-366391/47.
DR N-PSDB; AAT04248.
XX
PT Modified Cryptomeria japonica (Cry j) I peptide(s) - useful for
PT treating allergy to Japanese cedar pollen allergen or
PT immunologically cross reactive allergens
XX
PS Disclosure; Figure 1; 60pp; English.
XX
XX Novel peptides of cry j I have been modified as a part of a
XX preformulation scheme to develop an optimised drug product for
XX therapeutic treatment of humans suffering from allergy to Japanese
XX cedar pollen allergen or an allergen which is immunologically cross
XX reactive with Japanese cedar pollen allergen. Such modified peptides
XX possess certain characteristics which render them particularly
XX suitable for drug product formulation. Peptide fragments of Cry j I,
XX modified and unmodified, are given in AAR82491-R82525.
XX
SQ Sequence 374 AA;
XX
Query Match 100.0%; Score 83; DB 16; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 KMPMYIAGYKTFDGR 15
 DB 102 KMPMYIAGYKTFDGR 116

RESULT 14

ID AAY25665 standard; protein; 374 AA.

AC AAY25665;

DT 30-SEP-1999 (first entry)

DE Cedar allergen 493632 Cry j IA protein fragment.

OS Major histocompatibility complex; class II; desensitizing; human;
 PN allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 PD chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 XX screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 XX cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 XX mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.

PN WO9934826-A1.

PD 15-JUL-1999.

PE 11-JAN-1999; 99WO-GB00080.

PR 21-SEP-1998; 98GB-0020474.

PR 09-JAN-1998; 98GB-0000445.

PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

PI Kay AB, Larche M;

DR WPI; 1999-458255/38.

PT Desensitizing patients to polypeptide allergens

PS Example 6; Page 73; 117pp; English.

XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC tenbrion molly beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a cedar (Cedrus sp.) allergen 493632 Cry j IA.

XX Sequence 374 AA;

Query Match 100.0%; Score 83; DB 20; Length 374;

Best Local Similarity 100.0%; Pred. No. 1.3e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTFDGR 15

DB 102 KMPMYIAGYKTFDGR 116

RESULT 15

AAY25669

ID AAY25669 standard; protein; 374 AA.

AC AAY25669;

DT 30-SEP-1999 (first entry)

DE Japanese cedar allergen 541802 Cry j I precursor protein fragment.

OS Major histocompatibility complex; class II; desensitizing; human;
 PN allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 PD chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 XX screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 XX cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 XX mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.

PN WO9934826-A1.

PD 15-JUL-1999.

PE 11-JAN-1999; 99WO-GB00080.

PR 21-SEP-1998; 98GB-0020474.

PR 09-JAN-1998; 98GB-0000445.

PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

PI Kay AB, Larche M;

DR WPI; 1999-458255/38.

PT Desensitizing patients to polypeptide allergens

PS Example 6; Page 75; 117pp; English.

XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC tenbrion molly beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 541802 Cry j I
 CC precursor.

XX Sequence 374 AA;

Query Match 100.0%; Score 83; DB 20; Length 374;

Best Local Similarity 100.0%; Pred. No. 1.3e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KMPMYIAGYKTFDGR 15

DB 102 KMPMYIAGYKTFDGR 116

Search completed: April 20, 2003, 13:06:07
 Job time: 19.1974 secs



KW sugi-pollinosis; allergic reaction; pollen.
 CC Synthetic.
 OS
 XX JF10259198-A.
 PN
 XX 29-SEP-1998.
 PD
 XX 22-DEC-1997; 97JP-0353448.
 PF
 XX 24-DEC-1996; 96JP-0343441.
 PR
 XX
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 DR WPI; 1998-577037/49.
 XX
 XX A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollinosis
 XX
 XX Claim 7; Page 18; 21pp; Japanese.
 PS
 CC AAM80339-58 represent epitopes for T cells, derived from the sugi
 CC allergen proteins Cryj1 (AAM80339-44, AAM80350-53 and AAM80356-58) and
 CC Cryj2 (AAM80345-49 and AAM80354-55). The peptides are useful for the
 CC treatment of sugi-pollinosis, an allergic reaction of the body to
 CC pollen.
 CC
 XX Sequence 17 AA;
 SQ
 Query Match 100.0%; Score 81; DB 19; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VDGIIAAYONPASWK 15
 |||||
 DB 2 VDGIIAAYONPASWK 16
 RESULT 5
 AAM27369
 ID AAM27369 standard; peptide; 80 AA.
 XX
 AC AAM27369;
 XX
 DT 24-MAR-1998 (first entry)
 XX
 DE Multi-epitope peptide used as immunotherapeutic agent #1.
 XX
 KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 OS Synthetic.
 OS
 XX WO9732600-A1.
 PN
 XX 12-SEP-1997.
 PD
 XX 10-MAR-1997; 97WO-JP00740.
 PF
 XX 10-MAR-1996; 96JP-0080702.
 PR
 XX (MEIP) MEIJI MILK PROD CO LTD.
 PA
 XX Dairiki K, Iwama A, Kino K, Kume A, Sone T;
 PI WPI; 1997-470495/43.
 DR
 XX Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens
 XX
 PS Claim 6; Page 31; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX
 XX Sequence 80 AA;
 SQ
 Query Match 100.0%; Score 81; DB 18; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VDGIIAAYONPASWK 15
 |||||
 DB 66 VDGIIAAYONPASWK 80
 RESULT 6
 AAM27370
 ID AAM27370 standard; peptide; 105 AA.
 XX
 AC AAM27370;
 XX
 DT 24-MAR-1998 (first entry)
 XX
 DE Multi-epitope peptide used as immunotherapeutic agent #2.
 XX
 KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 OS Synthetic.
 OS
 XX WO9732600-A1.
 PN
 XX 12-SEP-1997.
 PD
 XX 10-MAR-1997; 97WO-JP00740.
 PF
 XX 10-MAR-1996; 96JP-0080702.
 PR
 XX (MEIP) MEIJI MILK PROD CO LTD.
 PA
 XX Dairiki K, Iwama A, Kino K, Kume A, Sone T;
 PI WPI; 1997-470495/43.
 DR
 XX Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens
 XX
 PS Claim 6; Page 31; 58pp; Japanese.
 XX
 CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX
 XX Sequence 105 AA;
 SQ
 Query Match 100.0%; Score 81; DB 18; Length 105;
 Best Local Similarity 100.0%; Pred. No. 1.6e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGIIAAYONPASWK 15
 |||||
 DB 91 VDGIIAAYONPASWK 105

RESULT 7
 AAM27371

ID AAM27371 standard; peptide; 134 AA.

AC AAM27371;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #3.

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;

KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN WO9732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;

DR WPI; 1997-470495/43.

PT Peptide immuno-therapeutic agent to treat allergic diseases -

PT contains multi-epitope peptide containing T cell epitope regions

PT from different allergens

PS Claim 6; Page 32; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as

CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2

CC or more different allergens (preferably linked via arginine or lysine

CC dimers), where the T cell epitope regions have a positivity index

CC greater than 100 as measured in a patient group responding to the

CC allergen; have at least 70% reactivity with lymphocytes from patients

CC responding to the allergen; and are not reactive with immunoglobulin E

CC (IgE) antibodies from patients responsive to the allergen. The agent can

CC be used to prevent and treat a wide variety of allergic diseases, e.g. by

CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX Sequence 134 AA;

QY Query Match 100.0%; Score 81; DB 18; Length 134;

Best Local Similarity 100.0%; Pred. No. 2.1e-06; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGIIAAYONPASWK 15
 |||||
 DB 120 VDGIIAAYONPASWK 134

RESULT 8
 AAR69791

ID AAR69791 standard; Protein; 460 AA.

AC AAR69791;

DT 27-SEP-1995 (first entry)

DE Japonicum allergen residues 55-514.

XX Japonicum allergen; residues 55-514; induced histamine release;

KW anti-allergic peptide; IgE cross-linking inhibition.

OS Japonicum sp.

PN WO9502412-A.

PD 26-JAN-1995.

PF 15-JUL-1994; 94WO-JP01164.

PR 16-JUL-1993; 93JP-0177008.

PR 01-SEP-1993; 93JP-0217725.

PR 07-APR-1994; 94JP-0069336.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Kino K, Kohno Y, Komiyama N, Sone T;

DR WPI; 1995-067159/09.

DR N-PSDB; AAQ84044.

PT Peptide anti-allergic agent - inhibits cross-linking of allergen

PT with IgE antibody

PS Disclosure; Pages 26-27; 46pp; Japanese.

XX AAQ84044 encodes AAR69791 Japonicum allergen residues 55-514, from

CC which the anti-allergic peptides AAR69845-R69809 were derived.

CC The peptides ability to inhibit the cross-linking of an allergen,

CC to an IgE antibody can be used in the prevention and treatment of

CC allergic diseases.

XX Sequence 460 AA;

QY Query Match 100.0%; Score 81; DB 16; Length 460;

Best Local Similarity 100.0%; Pred. No. 9.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGIIAAYONPASWK 15
 |||||
 DB 66 VDGIIAAYONPASWK 80

RESULT 9
 AAR53690

ID AAR53690 standard; Protein; 514 AA.

AC AAR53690;

DT 01-FEB-1995 (first entry)

DE Japanese cedar pollen allergen Cry j II.

XX Cedar pollinosis; diagnostic.

XX Cryptomeria japonica.

PN WO9411512-A.

PD 26-MAY-1994.

PF 12-NOV-1993; 93WO-US11000.

PR 12-NOV-1992; 92US-0975179.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Brauer A, Kuo M, Pollock J, Yeung S;

DR WPI; 1994-183513/22.

DR N-PSDB; AAQ66048.

XX Allergic Cry j II protein and fragments from Japanese cedar
PT pollen - used to diagnose, treat and prevent Japanese cedar
PT pollinosis
XX
PS Claim 2; Fig 4; 89pp; English.
XX
CC The sequence is of a Japanese cedar pollen allergen Cry j
CC II. The protein and its fragments can be used for diagnosis and
CC treatment of Japanese cedar pollinosis and to identify similar
CC sequences in other plants.
CC See also AAR53692-6.
XX
SQ Sequence 514 AA;
Query Match 100.0%; Score 81; DB 15; Length 514;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDGIIAAAYQNPASWK 15
DB 120 VDGIIAAAYQNPASWK 134
RESULT 10
AAR74333 standard; Protein; 514 AA.
XX
AC AAR74333;
XX
DT 01-NOV-1995 (first entry)
XX
DE Japanese cedar pollen allergen.
XX
KM Japanese cedar; pollen; allergen; allergy; therapy; diagnostic;
KM desensitizer.
XX
OS Cryptomeria japonica.
XX
PN EP65500-A.
XX
PD 31-MAY-1995.
XX
PF 03-NOV-1994; 94EP-0308117.
XX
PR 27-DEC-1993; 93JP-0346814.
PR 05-NOV-1993; 93JP-0299151.
PR 20-DEC-1993; 93JP-0344596.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Kurimoto M, Namba M, Torigoe K;
PI WPI; 1995-195588/26.
DR N-PSDB; AAO90156.
XX
XX New Japanese cedar pollen allergen polypeptide - and DNA coding for it,
PT useful for treatment and diagnosis of cedar pollen allergy
XX
XX Claim 5; Page 26-28; 41pp; English.
XX
CC The gene encoding an allergen of Japanese cedar pollen was isolated
CC by PCR amplification using primers based on portions of the allergen
CC protein. The gene was used for recombinant allergen production in
CC E. coli (vector plasmid pKK-223-3).
XX
SQ Sequence 514 AA;
Query Match 100.0%; Score 81; DB 16; Length 514;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDGIIAAAYQNPASWK 15

DB 120 VDGIIAAAYQNPASWK 134
RESULT 11
AAR69792 standard; Protein; 514 AA.
XX
AC AAR69792;
XX
DT 27-SEP-1995 (first entry)
XX
DE Japonicum allergen.
XX
KM Japonicum allergen; induced histamine release; antiallergic peptide;
KM IGE cross-linking inhibition.
XX
OS Japonicum sp.
XX
PN WO9502412-A.
XX
PD 26-JAN-1995.
XX
PF 15-JUL-1994; 94WO-JP01164.
XX
PR 16-JUL-1993; 93JP-0177008.
PR 01-SEP-1993; 93JP-0217725.
PR 07-APR-1994; 94JP-0069336.
XX
PA (MEIP) MEIJI MILK PROD CO LTD.
XX
PI Kino K, Kohno Y, Komiyama N, Sone T;
XX
DR WPI; 1995-067159/09.
DR N-PSDB; AAO84045; AAO84046.
XX
PT Peptide antiallergic agent - inhibits cross-linking of allergen
PT with IGE antibody
XX
PS Example 3; Pages 27-28; 46pp; Japanese.
XX
CC AAO84045 encodes AAR69792 Japonicum allergen, from which the
CC antiallergic peptides AAR69845-R69809 were derived. The peptides ability
CC to inhibit the cross-linking of an allergen, to an IGE antibody can be
CC used in the prevention and treatment of allergic diseases.
XX
SQ Sequence 514 AA;
Query Match 100.0%; Score 81; DB 16; Length 514;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDGIIAAAYQNPASWK 15
DB 120 VDGIIAAAYQNPASWK 134
RESULT 12
AAR93599 standard; Protein; 514 AA.
XX
AC AAR93599;
XX
DT 16-AUG-1996 (first entry)
XX
DE Japan cedar pollen Cry j II allergen.
XX
KM Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;
KM Sugi pollinosis; diagnosis; treatment.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers

```

FT Peptide 1.54
FT /label= sig_peptide
FT Protein 55.514
FT /*label= mac_protein
XX
XX JP08047392-A.
XX
XX 20-FEB-1996.
XX
XX 07-NOV-1994; 94UP-0297840.
XX
XX 26-MAY-1994; 94UP-0134868.
XX
XX 05-NOV-1993; 93UP-0276773.
XX
XX (MEIP ) MEIJI MILK PROD CO LTD.
XX
XX WPI; 1996-166249/17.
XX
XX N-PSDB; AAT18102.
XX
XX Japan cedar pollen allergen Cry j II epitope - comprises at least
XX part of specified 460 amino acid protein
XX
XX Claim 1; Page 10-11; 17pp; Japanese.
XX
XX AAR93599 is a Japan cedar pollen Cry j II allergen which is useful
XX in the diagnosis, prevention and treatment of Sugi pollinosis,
XX the allergic reaction to Japan cedar pollen. Significant regions of
XX the allergen were identified using overlapping peptides of the full
XX epitope derived from a Cry j II antigen-specific T cell line
XX (see AAR97871-R97960). Amino acids 66-80 (AAR97884) and 186-200
XX (R978908) of the full mature 460 amino acid allergen are the most
XX allergenic of the 90 peptides tested.
XX
XX Sequence 514 AA;
XX
XX Query Match 100.0%; Score 81; DB 17; Length 514;
XX Best Local Similarity 100.0%; Pred. No. 1e-05;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 VDGIIAAYQNPASWK 15
XX |||||
XX DB 120 VDGIIAAYQNPASWK 134
XX
XX RESULT 13
XX AAR81586
XX ID AAR81586 standard; Protein; 514 AA.
XX
XX AAR81586;
XX
XX 24-MAY-1996 (first entry)
XX
XX Cedar pollen allergen A.
XX
XX Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
XX antibody; pollinosis; therapy; immunotherapy.
XX
XX Cryptomeria japonica.
XX
XX EP700929-A2.
XX
XX 13-MAR-1996.
XX
XX 08-SEP-1995; 95EP-0306295.
XX
XX 14-JUL-1995; 95JP-0200221.
XX
XX 10-SEP-1994; 94JP-0242137.
XX
XX 14-JUL-1995; 95JP-0200204.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Hino K, Saito S, Taniguchi Y;
XX

```

```

DR WPI; 1996-140976/15.
XX
XX New peptide(s) derived from cedar pollen allergens - activate
XX allergen-specific T-cells, but not allergen-specific IgE antibodies,
XX used for treating cedar pollinosis
XX
XX Claim 5; Page 29-30; 36pp; English.
XX
XX Synthetic peptides based on portions of cedar pollen allergens A
XX (AAR81586) and B (AAR81587) were tested for their ability to activate
XX cedar allergen-specific T-cells, but not allergen-specific IgE
XX antibodies. 6 peptides (AAR81580-R81585) were identified as T-cell
XX epitopes. These peptides, plus subsequences (AAR81573-79) essential
XX for T-cell recognition, and homologous peptides (AAR81588-96) can
XX be used as immunotherapeutic agents to treat or prevent cedar
XX pollinosis, avoiding side-effects such as anaphylaxis.
XX
XX Sequence 514 AA;
XX
XX Query Match 100.0%; Score 81; DB 17; Length 514;
XX Best Local Similarity 100.0%; Pred. No. 1e-05;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 VDGIIAAYQNPASWK 15
XX |||||
XX DB 120 VDGIIAAYQNPASWK 134
XX
XX RESULT 14
XX AAY25666
XX ID AAY25666 standard; protein; 514 AA.
XX
XX AAY25666;
XX
XX 30-SEP-1999 (first entry)
XX
XX Japanese cedar allergen 1076242 Cry j II precursor protein fragment.
XX
XX Major histocompatibility complex; class II; desensitizing; human;
XX allergen; grass; tree; weed; pollen; fungi; mould; food; insect; string;
XX chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX
XX Cedrus sp.
XX
XX WO9934826-A1.
XX
XX 15-JUL-1999.
XX
XX 11-JAN-1999; 99WO-GB00080.
XX
XX 21-SEP-1998; 98GB-0020474.
XX
XX 09-JAN-1998; 98GB-0000445.
XX
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
XX Kay AB, Larche M;
XX
XX WPI; 1999-458255/38.
XX
XX Desensitizing patients to polypeptide allergens
XX
XX Example 6; Page 74; 117pp; English.
XX
XX This invention describes a novel method of desensitizing a patient to a
XX polypeptide allergen and comprises administering to the patient a peptide
XX derived from the allergen where restriction to a MHC class II molecule
XX possessed by the patient can be demonstrated for the peptide and the
XX peptide is able to induce a late phase response in an individual who
XX possesses the MHC class II molecule. The methods can be used for
XX desensitizing patients to allergens present in e.g. grass, tree and weed
XX

```

CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076242 Cys j II
 CC precursor.

XX
 XX Sequence 514 AA;
 SQ

Query Match 100.0%; Score 81; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGIIAAYQNPAWSK 15
 |||||
 DB 120 VDGIIAAYQNPAWSK 134

RESULT 15

AAV25667
 ID AAV25667 standard; protein; 514 AA.

XX AAV25667;

DT 30-SEP-1999 (first entry)

DE Japanese cedar allergen 1076241 Cys j II protein fragment.

XX Major histocompatibility complex; class II; desensitizing; human;
 KW allergen; grass; tree; weed; pollen; mould; food; insect; sting;
 KW chironomidae; spider; mice; housefly; fruit fly; sheep blow fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

XX Cedrus sp.

XX WO9934826-A1.

XX 15-JUL-1999.

XX 11-JAN-1999; 99WO-GB00080.

XX 21-SEP-1998; 98GB-0020474.

XX 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Kay AB, Larche M;

XX WPI; 1999-458255/38.

XX Desensitizing patients to polypeptide allergens

XX Example 6; Page 74; 117p; English.

XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to

CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 1076241 Cys j II.

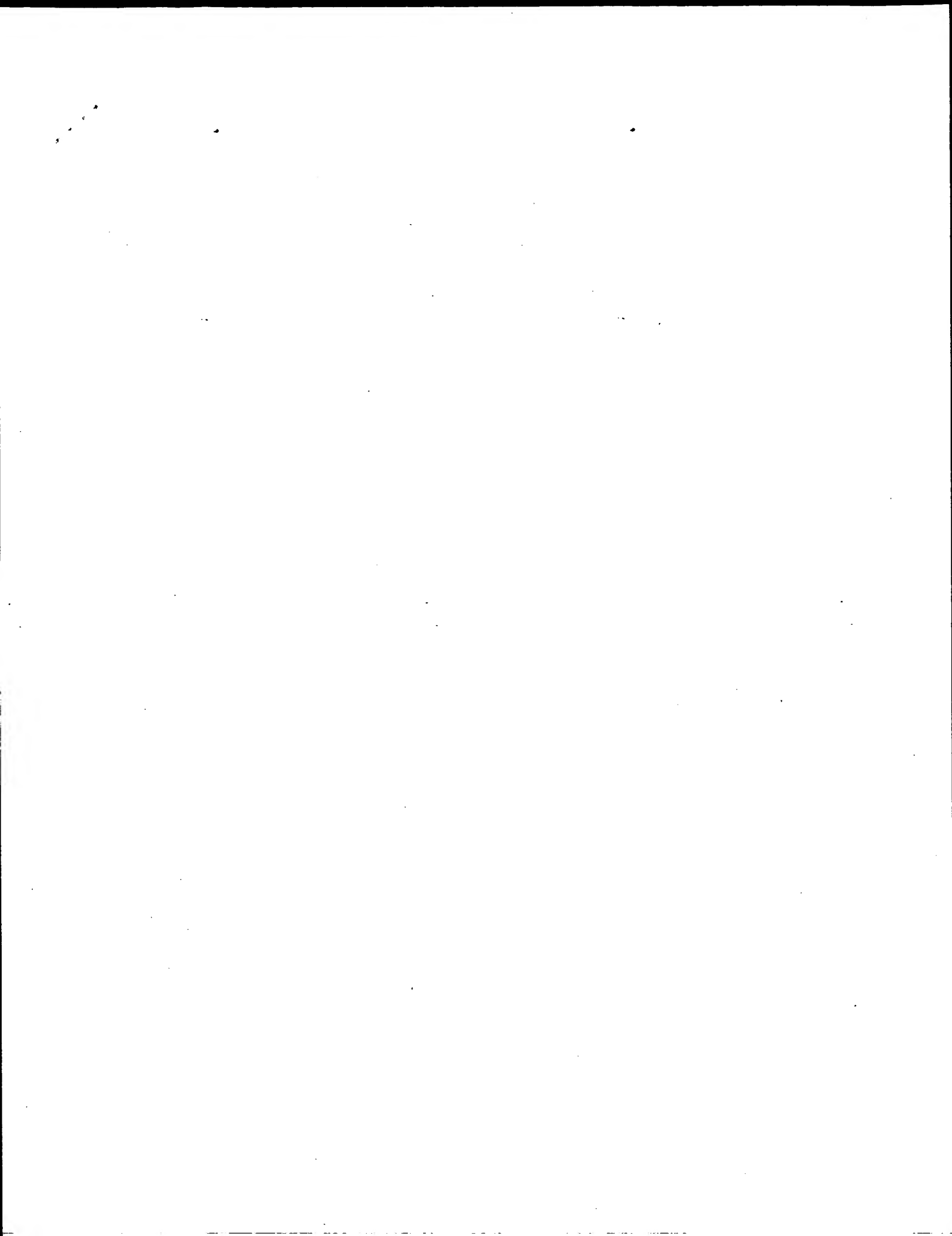
XX
 XX Sequence 514 AA;
 SQ

Query Match 100.0%; Score 81; DB 20; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDGIIAAYQNPAWSK 15
 |||||
 DB 120 VDGIIAAYQNPAWSK 134

Search completed: April 20, 2003, 13:06:09

Job time : 19.1974 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 18.1974 Seconds
(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524D-36
Perfect score: 80
Sequence: 1 PCVFKRVSNVITHG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	15	16	AAAR89291
2	80	100.0	20	15	AAAR45552
3	80	100.0	20	16	AAAR82501
4	80	100.0	30	19	AAW44683
5	80	100.0	353	16	AAAR75388
6	80	100.0	353	17	AAAR81587
7	80	100.0	373	20	AAAY25664
8	80	100.0	374	20	AAAY25668
9	80	100.0	374	14	AAAR31937
10	80	100.0	374	15	AAAR45541

11	80	100.0	374	15	AAAR60166
12	80	100.0	374	16	AAAR82490
13	80	100.0	374	20	AAAY25665
14	80	100.0	374	20	AAAY25669
15	73	91.2	15	19	AAW57750
16	64	80.0	13	22	AAAB84112
17	64	80.0	36	22	AAAB84102
18	64	80.0	80	18	AAAM27369
19	64	80.0	105	18	AAAM27370
20	64	80.0	134	18	AAAM27371
21	61	76.2	370	15	AAAR45578
22	58	72.5	20	19	AAW42133
23	58	72.5	354	17	AAW04344
24	58	72.5	354	19	AAW42121
25	58	72.5	375	17	AAW04345
26	57	71.2	264	21	AAAG39243
27	57	71.2	270	21	AAAG39243
28	57	71.2	274	21	AAAG39242
29	57	71.2	280	21	AAAG39242
30	57	71.2	384	21	AAAG39241
31	57	71.2	384	23	ABR90972
32	57	71.2	390	21	AAAG29241
33	56	70.0	36	21	AAAG27996
34	56	70.0	45	21	AAAG27995
35	56	70.0	266	21	AAAG39368
36	56	70.0	272	21	AAAG42704
37	56	70.0	274	21	AAAG32458
38	56	70.0	276	21	AAAG39367
39	56	70.0	349	21	AAAG32457
40	56	70.0	394	21	AAAG32456
41	56	70.0	394	23	ABR93044
42	56	70.0	394	23	ABR93045
43	56	70.0	404	21	AAAG42703
44	56	70.0	404	23	ABR93086
45	56	70.0	406	21	AAAG06909

ALIGNMENTS

RESULT 1
ID AAR89291 standard; peptide; 15 AA.
XX AAR89291;
AC
XX 12-MAR-1996 (first entry)
DT
XX Japanese cedar pollen allergen Cryj I derived T-cell epitope peptide.
DE
XX Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides;
KW prevention; treatment; cryptomeria pollenosis.
XX
XX Cryptomeria japonica.
OS
XX JP07118295-A.
PN
XX 09-MAY-1995.
PD
XX 20-OCT-1993; 93JP-0262626.
PE
XX 20-OCT-1993; 93JP-0262626.
PR
XX 20-OCT-1993; 93JP-0262626.
XX
XX (MEIP) MEIJI MILK PROD CO LTD.
PA
XX WPI; 1995-203834/27.
DR
XX
XX New cryptomeria pollen allergen T-cell epitope peptide - used for
PT prevention, treatment and investigation of Japanese cedar pollenosis
XX
XX Claim 5; Page 2; 8pp; Japanese.
PS
XX AAR75388 is the Japanese cedar pollen allergen Cryj I, from which the

CC T-cell epitope peptides AAR8289-R89295 were derived. The peptides
CC can be used for the prevention and treatment of Cryptomeria
CC pollenosis, and also for the investigation of pollenosis.

SQ Sequence 15 AA;

Query Match 100.0%; Score 80; DB 16; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIRKRVSNVIHG 15
DB 1 PCVFIRKRVSNVIHG 15

RESULT 2
AAR45552
ID AAR45552 standard; Protein; 20 AA.

AC AAR45552;

DT 13-JUN-1994 (first entry)

DE Cry j I pollen allergen peptide CJI-11.

KM Japanese cedar; detection; allergy; treatment; diagnosis;

KW T cell epitope; sensitivity.

OS Cryptomeria japonica.

XX WO9401560-A.

XX 20-JAN-1994.

XX 15-JAN-1993; 93WO-US00139.

XX 01-SEP-1992; 92US-0938990.

XX 10-JUL-1992; 92WO-US05661.

XX (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Bond UF, Garman RD, Griffith IJ, Kuo M, Pollock J;

XX WPI; 1994-035066/04.

PT Antigens derived from Japanese cedar pollen allergen Cry j I -

PT contain at least two T cell epitope(s), used to treat or diagnose

XX allergy

XX Claim 1; Fig 13; 137pp; English.

CC The sequence is that of an isolated peptide of the Japanese cedar

CC pollen allergen Cry j I (amino acids 101-120). The peptide, CJI-11,

CC can be used for the treatment and diagnosis of allergies associated

CC with Japanese cedar pollen. It has enhanced therapeutic properties

CC but reduced side effects compared to naturally occurring allergens.

XX SQ Sequence 20 AA;

OY 1 PCVFIRKRVSNVIHG 15

DB 6 PCVFIRKRVSNVIHG 20

RESULT 3

AAR82501
ID AAR82501 standard; Protein; 20 AA.

AC AAR82501;

XX 15-APR-1996 (first entry)

DT Cry j I Japanese Cedar pollen allergen peptide fragment (CJI-11).

DE Cry j I Japanese cedar pollen allergen; modified; drug production;

XX allergy; Cryptomeria japonica.

OS Cryptomeria japonica.

XX WO9527786-A1.

XX 19-OCT-1995.

XX 06-APR-1995; 95WO-US04249.

XX 06-DEC-1994; 94US-0350225.

XX 08-APR-1994; 94US-0226248.

XX (IMMU-) IMMUNOLOGIC PHARM CORP.

XX Chen X, Evans S, Franzen HM, Kuo M, Powers SP;

XX Shaked Z;

XX WPI; 1995-366391/47.

XX Modified Cryptomeria japonica (Cry j) I peptide(s) - useful for

XX treating allergy to Japanese cedar pollen allergen or

XX immunologically cross reactive allergens

XX Disclosure; Figure 2; 60pp; English.

XX Novel peptides of cry j I have been modified as a part of a

XX preformulation scheme to develop an optimised drug product for

XX therapeutic treatment of humans suffering from allergy to Japanese

XX cedar pollen allergen or an allergen which is immunologically cross

XX reactive with Japanese cedar pollen allergen. Such modified peptides

XX possess certain characteristics which render them particularly

XX suitable for drug product formulation. Peptide fragments of Cry j I,

XX modified and unmodified, are given in AAR82491-R82525. This peptide

XX fragment corresponds to amino acids 101-120 of the allergen mature

XX protein.

XX SQ Sequence 20 AA;

OY 1 PCVFIRKRVSNVIHG 15

DB 6 PCVFIRKRVSNVIHG 20

RESULT 4

AAM44683
ID AAM44683 standard; peptide; 30 AA.

AC AAM44683;

DT 01-MAY-1998 (first entry)

DE T-cell epitope peptide #2 of sugi pollen antigen.

XX T-cell epitope; sugi pollen antigen; sugi pollinosis.

XX Synthetic.

XX Cryptomeria japonica.

XX JPI0007700-A.

XX 13-JAN-1998.

PF 24-JUN-1996; 96JP-0163287.
 XX
 PT 24-JUN-1996; 96JP-0163287.
 XX
 XX (DAIL) DAICEL CHEM IND LTD.
 PA (MEIJ) MEIJI SEIKA KAISHA LTD.
 XX
 DR WPI; 1996-133630/13.
 XX
 PT T cell epitope peptide of sugi pollen antigen - useful in the
 XX treatment of sugi pollinosis
 PS Claim 1; Page 4; 14pp; Japanese.
 XX
 CC T-cell epitope peptides AAR44682-88 and their derivatives react with
 CC sugi pollinosis patient peripheral blood T lymphocytes. A composition
 CC prepared by combining at least 2 of the above peptides and/or their
 CC derivatives is used for the prevention and treatment of sugi
 CC pollinosis.
 CC
 SQ Sequence 30 AA;

Query Match 100.0%; Score 80; DB 19; Length 30;
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIKRVSNVTHG 15
 DB 6 PCVFIKRVSNVTHG 20

RESULT 5
 AAR75388 standard; protein; 353 AA.

AC AAR75388;

DT 12-MAR-1996 (first entry)

DE Japanese cedar pollen allergen Cry I.

KW Japanese cedar pollen allergen; Cry I; T-cell epitope; peptides;
 XX prevention; treatment; cryptomeria pollinosis.

OS Cryptomeria japonica.

Key Location/Qualifiers

FT Peptide

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

FT /note= "T-cell epitope peptide"

XX
 PT New cryptomeria pollen allergen T-cell epitope peptide - used for
 PT prevention, treatment and investigation of Japanese cedar pollinosis
 XX
 XX Disclosure; Figs 1-2; 8pp; Japanese.
 PS
 CC AAR75388 is the Japanese cedar pollen allergen Cry I, from which the
 CC T-cell epitope peptides AAR8289-88295 were derived. The peptides
 CC can be used for the prevention and treatment of cryptomeria
 CC pollinosis, and also for the investigation of pollinosis.
 CC
 SQ Sequence 353 AA;

Query Match 100.0%; Score 80; DB 16; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PCVFIKRVSNVTHG 15
 DB 106 PCVFIKRVSNVTHG 120

RESULT 6
 AAR1587
 ID AAR81587 standard; protein; 353 AA.

AC AAR81587;

DT 24-MAY-1996 (first entry)

DE Cedar pollen allergen B.

KW Cedar pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
 XX antibody; pollinosis; therapy; immunotherapy.

OS Cryptomeria japonica.

PN EP700929-A2.

PD 13-MAR-1996.

PF 08-SEP-1995; 95EP-0306295.

PR 14-JUL-1995; 95JP-0200221.

PR 10-SEP-1994; 94JP-0242137.

PR 14-JUL-1995; 95JP-0200204.

XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

XX Hino K, Saito S, Taniguchi Y;

XX WPI; 1996-140976/15.

DR

PT New peptide(s) derived from cedar pollen allergens - activate

PT allergen-specific T-cells, but not allergen-specific IgE antibodies,

PT used for treating cedar pollinosis

PS Claim 5; Page 31-32; 36pp; English.

XX

XX Synthetic peptides based on portions of cedar pollen allergens A

XX (AAR1586) and B (AAR1587) were tested for their ability to activate

XX cedar allergen-specific T-cells, but not allergen-specific IgE

XX antibodies. 6 Peptides (AAR1580-R1585) were identified as T-cell

XX epitopes. These peptides, plus subsequences (AAR1571-79) essential

XX for T-cell recognition, and homologous peptides (AAR1578-96) can

XX be used as immunotherapeutic agents to treat or prevent cedar

XX pollinosis, avoiding side-effects such as anaphylaxis.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 100.0%; Score 80; DB 17; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCVFIRKRVSNVIING 15
 DB 106 PCVFIRKRVSNVIING 120

RESULT 7

AAV25664 standard; protein; 373 AA.

AAV25664;

30-SEP-1999 (first entry)

Cedar allergen 493634 Cry j IB protein fragment.

Major histocompatibility complex; class II; desensitizing; human;
 allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 mite; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

Cedrus sp.

WO9934826-A1.

15-JUL-1999.

11-JAN-1999; 99WO-GB00080.

21-SEP-1998; 98GB-0020474.

09-JAN-1998; 98GB-0000445.

(IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

Kay AB, Larche M;

WPI; 1999-458255/38.

Desensitizing patients to polypeptide allergens

Example 6; Page 73; 117pp; English.

This invention describes a novel method of desensitizing a patient to a polypeptide allergen and comprises administering to the patient a peptide derived from the allergen where restriction to a MHC Class II molecule possessed by the patient can be demonstrated for the peptide and the peptide is able to induce a late phase response in an individual who possesses the MHC Class II molecule. The methods can be used for desensitizing patients to allergens present in e.g. grass, tree and weed (including ragweed) pollens, fungi and moulds, foods, stinging insects, the chironomidae (non-biting midges), spiders and mites, housefly, fruit fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee, non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig, sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to produce immunological vaccines which may be used to prevent and/or treat conditions involving hypersensitivity to allergens. This sequence represents a cedar (Cedrus sp.) allergen 493634 Cry j IB.

Sequence 373 AA;

Query Match 100.0%; Score 80; DB 20; Length 373;

Best Local Similarity 100.0%; Pred. No. 5.4e-06; Mismatches 0; Indels 0; Gaps 0;

1 PCVFIRKRVSNVIING 15

126 PCVFIRKRVSNVIING 140

RESULT 8

AAV25668 standard; protein; 373 AA.

AAV25668;

30-SEP-1999 (first entry)

Japanese cedar allergen 541803 Cry j I precursor protein fragment.

Major histocompatibility complex; class II; desensitizing; human;
 allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 mite; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

Cedrus sp.

WO9934826-A1.

15-JUL-1999.

11-JAN-1999; 99WO-GB00080.

21-SEP-1998; 98GB-0020474.

09-JAN-1998; 98GB-0000445.

(IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

Kay AB, Larche M;

WPI; 1999-458255/38.

Desensitizing patients to polypeptide allergens

Example 6; Page 75; 117pp; English.

This invention describes a novel method of desensitizing a patient to a polypeptide allergen and comprises administering to the patient a peptide derived from the allergen where restriction to a MHC Class II molecule possessed by the patient can be demonstrated for the peptide and the peptide is able to induce a late phase response in an individual who possesses the MHC Class II molecule. The methods can be used for desensitizing patients to allergens present in e.g. grass, tree and weed (including ragweed) pollens, fungi and moulds, foods, stinging insects, the chironomidae (non-biting midges), spiders and mites, housefly, fruit fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee, non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig, sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to produce immunological vaccines which may be used to prevent and/or treat conditions involving hypersensitivity to allergens. This sequence represents a Japanese cedar (Cedrus sp.) allergen 541803 Cry j I precursor.

Sequence 373 AA;

Query Match 100.0%; Score 80; DB 20; Length 373;

Best Local Similarity 100.0%; Pred. No. 5.4e-06; Mismatches 0; Indels 0; Gaps 0;

1 PCVFIRKRVSNVIING 15

126 PCVFIRKRVSNVIING 140

RESULT 9

AAV25668 standard; protein; 374 AA.

AAV25668;

03-JUN-1993 (first entry)


```

XX      ?
DE      Cry j I.
XX
XX      Japanese cedar pollen; allergen; antigen; allergy; B cell; T cell.
XX
XX      Cryptomeria japonica.
XX
FH      Key      Location/Qualifiers
FT      Peptide  1..21
FT      Protein  /note= "signal peptide"
FT      Protein  22..374
FT      Protein  /note= "mature Cry j I"
XX
XX      WO9301213-A.
XX
XX      21-JAN-1993.
XX
XX      10-JUL-1992; 92WO-US05661.
XX
XX      12-JUL-1991; 91US-0729134.
XX      15-JUL-1991; 91US-0730452.
XX
XX      (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
XX      Bond JF, Griffith IU, Pollock J;
XX
XX      WPI; 1993-045434/05.
XX      N-PSDB; AAQ35304.
XX
XX      Nucleic acid sequence encoding Cryptomeria japonica allergen -
XX      for the diagnosis treatment and prevention of allergic reactions
XX      to Japanese cedar pollen
XX
XX      Claim 11; Page 42; 69pp; English.
XX
XX      Fresh pollen and staminate cone samples were collected from a single
XX      Cryptomeria japonica (Japanese cedar) tree. RNA was prep'd. and used
XX      to synthesize cDNA. The cDNA was subjected to successive rounds of
XX      PCR to yield a full length Cry j I clone. Cry j I or an antigenic
XX      fragment of it may be used for detecting, treating and preventing an
XX      allergic response to Japanese cedar pollen allergen. It is capable of
XX      modifying both the B and T cell response to Cry j I and T cell response
XX      to a Cry j I antigen.
XX
XX      Sequence 374 AA;
XX
XX      Query Match      100.0%; Score 80; DB 14; Length 374;
XX      Best Local Similarity 100.0%; Pred. No. 5.4e-06;
XX      Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY      1 PCVFIRKRVSNVTHG 15
DB      127 PCVFIRKRVSNVTHG 141

RESULT 10
AAR45541
ID      AAR45541 standard; Protein; 374 AA.
XX
XX      AAR45541;
XX
XX      13-JUL-1994 (first entry)
XX
XX      Cry j I pollen allergen.
XX
XX      Japanese cedar; detection; allergy; treatment; diagnosis;
XX      T cell epitope; sensitivity.
XX
XX      Cryptomeria japonica.
XX
XX      Key      Location/Qualifiers
XX      Peptide  1..21
XX      Protein  /note= "signal peptide"

```

```

FT      Peptide  22..374
FT      Protein  /note= "mature peptide"
XX
XX      WO9401560-A.
XX
XX      20-JAN-1994.
XX
XX      15-JAN-1993; 93WO-US00139.
XX
XX      01-SEP-1992; 92US-0938990.
XX      10-JUL-1992; 92WO-US05661.
XX
XX      (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
XX      Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;
XX
XX      WPI; 1994-035066/04.
XX      N-PSDB; AAQ55271.
XX
XX      Antigens derived from Japanese cedar pollen allergen Cry j I -
XX      contain at least two T cell epitope(s), used to treat or diagnose
XX      allergy
XX
XX      Disclosure; Fig 4; 137pp; English.
XX
XX      The sequence is that of the Japanese cedar pollen allergen
XX      Cry j I which contains at least two T cell epitopes. Peptide
XX      antigens derived from it can be used for the treatment and
XX      diagnosis of allergies associated with Japanese cedar pollen.
XX      The peptides have enhanced therapeutic properties but reduced
XX      side effects compared to naturally occurring allergens.
XX
XX      Sequence 374 AA;
XX
XX      Query Match      100.0%; Score 80; DB 15; Length 374;
XX      Best Local Similarity 100.0%; Pred. No. 5.4e-06;
XX      Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY      1 PCVFIRKRVSNVTHG 15
DB      127 PCVFIRKRVSNVTHG 141

RESULT 11
AAR60166
ID      AAR60166 standard; Protein; 374 AA.
XX
XX      AAR60166;
XX
XX      24-MAR-1995 (first entry)
XX
XX      Japanese cedar pollen antigen CryjI.
XX
XX      Japanese cedar; pollen antigen; allergen; CryjI; sugi; pollinosis.
XX
XX      Cryptomeria japonica.
XX
XX      Key      Location/Qualifiers
XX      Peptide  1..21
XX      Protein  /label= signal_peptide
XX      Protein  22..374
XX      Protein  /label= mature_CryjI
XX
XX      JP06197768-A.
XX
XX      19-JUL-1994.
XX
XX      07-JAN-1993; 93JP-0001116.
XX
XX      07-JAN-1993; 93JP-0001116.
XX
XX      (MEIJ) MEIJI SEIKA KAISHA.
XX

```

DR WPI, 1994-266860/33.
 DR N-PSDB; AAQ71601.
 XX
 PT Sugi (Japanese cedar) pollen antigen Cryj1 - is useful for
 diagnosis, treatment and prevention of sugi pollinosis
 XX
 PS Claim 2; Page 5-7; 9pp; Japanese.
 XX
 CC The coding sequence for the Japanese cedar ("sugi") pollen allergen
 CC Cryj1 was isolated from a cDNA library prepared from polyA mRNA. All
 CC or part of the Cryj1 protein can be used for diagnosis, treatment
 CC and prevention of sugi pollinosis.
 CC
 SQ Sequence 374 AA;
 Query Match 100.0%; Score 80; DB 15; Length 374;
 Best Local Similarity 100.0%; Pred. No. 5.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PCVFIKRVSNVTHG 15
 DB 127 PCVFIKRVSNVTHG 141
 RESULT 12
 AAR82490
 ID AAR82490 standard; Protein; 374 AA.
 XX
 AC AAR82490;
 XX
 DT 15-APR-1996 (first entry)
 XX
 DE Cry j I Japanese Cedar pollen allergen.
 XX
 KM Cry j I; Japanese cedar pollen allergen; modified; drug production;
 KM allergy; Crypomeria japonica.
 XX
 OS Crypomeria japonica.
 XX
 PN WO9527786-A1.
 XX
 PD 19-OCT-1995.
 XX
 PF 06-APR-1995; 95WO-US04249.
 XX
 PR 06-DEC-1994; 94US-0350225.
 PR 08-APR-1994; 94US-0226248.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP;
 PI Shaked Z;
 XX
 DR WPI; 1995-366391/47.
 DR N-PSDB; AAT04248.
 XX
 PT Modified Crypomeria japonica (Cry j) I peptide(s) - useful for
 PT treating allergy to Japanese cedar pollen allergen or
 PT immunologically cross reactive allergens
 XX
 PS Disclosure; Figure 1; 60pp; English.
 XX
 CC Novel peptides of cry j I have been modified as a part of a
 CC preformulation scheme to develop an optimised drug product for
 CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j I,
 CC modified and unmodified, are given in AAR82491-88255.
 CC
 SQ Sequence 374 AA;

Query Match 100.0%; Score 80; DB 16; Length 374;
 Best Local Similarity 100.0%; Pred. No. 5.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PCVFIKRVSNVTHG 15
 DB 127 PCVFIKRVSNVTHG 141
 RESULT 13
 AAY25665
 ID AAY25665 standard; protein; 374 AA.
 XX
 AC AAY25665;
 XX
 DT 30-SEP-1999 (first entry)
 XX
 DE Cedar allergen 493632 Cry j IA protein fragment.
 XX
 KM Major histocompatibility complex; class II; desensitizing; human;
 KM allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
 KM chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KM screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KM cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KM mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX
 OS Cedrus sp.
 XX
 PN WO934826-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 11-JAN-1999; 99WO-GB00080.
 XX
 PR 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Kay AB, Larche M;
 XX
 DR WPI, 1999-458255/38.
 XX
 PT Desensitizing patients to polypeptide allergens
 XX
 PS Example 6; Page 73; 117pp; English.
 XX
 CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a cedar (Cedrus sp.) allergen 493632 Cry j IA.
 CC
 SQ Sequence 374 AA;
 Query Match 100.0%; Score 80; DB 20; Length 374;
 Best Local Similarity 100.0%; Pred. No. 5.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PCVFIKRVSNVTHG 15
 DB 127 PCVFIKRVSNVTHG 141

```

RESULT 14
AAW57750
ID AAW57750 standard; protein; 374 AA.
XX
XX
AC AAW57750;
XX
XX
DT 30-SEP-1999 (first entry)
XX
XX
DE Japanese cedar allergen 541802 Cry j I precursor protein fragment.
XX
XX
Major histocompatibility complex; class II; desensitizing; human;
XX
XX
allergen; grass; tree; weed; pollen; fungi; mold; food; insect; sting;
XX
XX
chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX
XX
screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX
XX
cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX
XX
mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX
XX
Cedrus sp.
XX
XX
WO9934826-A1.
XX
XX
PD 15-JUL-1999.
XX
XX
PF 11-JAN-1999; 99WO-GB00080.
XX
XX
XX 21-SEP-1998; 98GB-0020474.
XX
XX
PR 09-JAN-1998; 98GB-0000445.
XX
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
XX
XX Kay AB, Larche M;
XX
XX
DR WPI; 1999-458255/38.
XX
XX
PT Desensitizing patients to polypeptide allergens
XX
XX
PS Example 6; Page 75; 117pp; English.
XX
XX
CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a Japanese cedar (Cedrus sp.) allergen 541802 Cry j I
CC precursor.
XX
XX
SQ Sequence 374 AA;
XX
XX
Query Match 100.0%; Score: 80; DB 20; Length 374;
XX
XX
Best Local Similarity 100.0%; Pred. No. 5.4e-06;
XX
XX
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
OY 1 PCVFIRKRVNVIHG 15
XX
XX
DB 127 PCVFIRKRVNVIHG 141
XX
XX
RESULT 15
AAW57750
ID AAW57750 standard; peptide; 15 AA.
XX

```

```

AC AAW57750;
XX
XX
DT 17-SEP-1998 (first entry)
XX
XX
DE Residues 106-120 of Cry j 1.
XX
XX
XX Cry j 1; Japanese cedar pollen antigen; allergy; immunotherapy;
XX
XX
XX HLA class II molecule.
XX
XX
OS Cryptococcosis japonica.
XX
XX
XX WO9820902-A1.
XX
XX
XX 22-MAY-1998.
XX
XX
XX 12-NOV-1997; 97WO-JP04129.
XX
XX
XX 13-NOV-1996; 96JP-0302053.
XX
XX
XX (MEIP ) MEIJI MILK PROD CO LTD.
XX
XX
XX Dairiki K, Kino K, Kume A, Sone T;
XX
XX
XX WPI; 1998-297617/26.
XX
XX
XX Peptides derived from Japanese cedar pollen antigens are
XX
XX
XX immunotherapeutic agents - useful for allergy treatment and typing
XX
XX
XX HLA class II molecules in allergy sufferers
XX
XX
XX Claim 12; Page 26; 50pp; Japanese.
XX
XX
XX This sequence represents residues 106-120 of the Cry j 1 protein, and
XX
XX
XX is a peptide of the invention. The peptides are derived from Japanese
XX
XX
XX cedar pollen antigens, and are used as immunotherapeutic agents in the
XX
XX
XX treatment of allergy. The peptides can be used for identification and
XX
XX
XX typing of the particular HLA class II molecules in an allergy sufferer,
XX
XX
XX and also for peptide immunotherapy of an allergy. Using these peptides
XX
XX
XX the immunotherapy can be targeted more specifically to the requirements
XX
XX
XX of the individual patient, allowing more effective treatment of an
XX
XX
XX allergy, including those patients for whom treatment with a conventional
XX
XX
XX immunotherapeutic agent is ineffective.
XX
XX
XX Sequence 15 AA;
XX
XX
XX Query Match 91.3%; Score 73; DB 19; Length 15;
XX
XX
XX Best Local Similarity 93.3%; Pred. No. 2.7e-06;
XX
XX
XX Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX
XX OY 1 PCVFIRKRVNVIHG 15
XX
XX
XX DB 1 PCVFIRKRVNVIHG 15
XX
XX
XX Search completed: April 20, 2003, 13:06:08
XX
XX
XX Job time : 19.1974 secs
XX

```



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05; Search time 18.1974 Seconds
(without alignments)
109.838 Million cell updates/sec

Title: US-09-142-524d-57

Perfect score: 78

Sequence: 1 KSMKVTVAENQFGPN 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq 101002:*

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	15	AA889293	Japanese cedar pol
2	78	100.0	15	AA857555	Residues 211-225 o
3	78	100.0	17	AAW80349	Sugi allergen prot
4	78	100.0	20	AA845563	Cry j I pollen all
5	78	100.0	20	AA82512	Cry j I Japanese C
6	78	100.0	20	AAW42144	T-cell epitope pep
7	78	100.0	26	AA845591	Cry j I pollen all
8	78	100.0	26	AA845592	Cry j I pollen all
9	78	100.0	26	AA845594	Cry j I pollen all
10	78	100.0	28	AA845589	Cry j I pollen all

Result No.	Score	Query Match	Length	ID	Description
11	78	100.0	28	AA845590	Cry j I pollen all
12	78	100.0	28	AA845593	Cry j I pollen all
13	78	100.0	30	AA845587	Cry j I pollen all
14	78	100.0	30	AA845588	Cry j I pollen all
15	78	100.0	30	AA845587	Cry j I pollen all
16	78	100.0	36	AA845586	Cry j I pollen all
17	78	100.0	35	AA875388	Cry j I pollen all
18	78	100.0	35	AA81587	Japanese cedar pol
19	78	100.0	35	AAW04344	Cedar pollen aller
20	78	100.0	35	AAW04344	Chamaecyparis obtu
21	78	100.0	36	AA845577	Japanese cyprus p
22	78	100.0	37	AA845577	Jun s I, Juniperu
23	78	100.0	37	AA845577	Cedar allergen 493
24	78	100.0	37	AA845577	Japanese cedar all
25	78	100.0	37	AA845577	Cry j I. Crypome
26	78	100.0	37	AA845577	Cry j I pollen all
27	78	100.0	37	AA845577	Japanese cedar pol
28	78	100.0	37	AA845577	Cry j I Japanese C
29	78	100.0	37	AA845577	Cedar allergen 493
30	78	100.0	37	AA845577	Japanese cedar all
31	78	100.0	37	AA845577	Chamaecyparis obtu
32	78	100.0	37	AA845577	Immunomodulatory p
33	78	100.0	37	AA845577	Artificial sequenc
34	78	100.0	37	AA845577	Artificial sequenc
35	78	100.0	37	AA845577	Artificial sequenc
36	78	100.0	37	AA845577	Artificial sequenc
37	78	100.0	37	AA845577	Cedar pollen aller
38	78	100.0	37	AA845577	Cedar pollen aller
39	78	100.0	37	AA845577	Cedar pollen aller
40	78	100.0	37	AA845577	Cedar pollen aller
41	78	100.0	37	AA845577	Artificial sequenc
42	78	100.0	37	AA845577	Artificial sequenc
43	78	100.0	37	AA845577	Cedar pollen aller
44	78	100.0	37	AA845577	Cedar pollen aller
45	78	100.0	37	AA845577	Amino acid sequenc

ALIGNMENTS

RESULT 1

AA889293 standard; peptide; 15 AA.

AA889293;

12-MAR-1996 (first entry)

Japanese cedar pollen allergen Cryj I derived T-cell epitope peptide.

Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides;

prevention; treatment; cryptomeria pollinosis.

Cryptomeria japonica.

JP07118295-A.

09-MAY-1995.

20-OCT-1993; 93JP-0262626.

20-OCT-1993; 93JP-0262626.

20-OCT-1993; 93JP-0262626.

(MEIP) MEIJI MILK PROD CO LTD.

WPI; 1995-203834/27.

New cryptomeria pollen allergen T-cell epitope peptide - used for

prevention, treatment and investigation of Japanese cedar pollinosis

Claim 5; Page 2; 8pp; Japanese.

AA875388 is the Japanese cedar pollen allergen Cryj I, from which the

CC T-cell epitope peptides AAR89289-R89295 were derived. The peptides
 CC can be used for the prevention and treatment of Cryptomeria
 CC pollenosis, and also for the investigation of pollenosis.

XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 78; DB 16; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.2e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15
 |||||
 DB 1 KSMKVTVAFNQFGPN 15

RESULT 2
 ID AAM57755
 XX AAM57755 standard; peptide; 15 AA.

AC AAM57755;

DT 17-SEP-1998 (first entry)

DE Residues 211-225 of Cry j 1.

KM Cry j 1; Japanese cedar pollen antigen; allergy; immunotherapy;
 KM HLA class II molecule.

OS Cryptomeria japonica.

PN WO9820902-A1.

PD 22-MAY-1998.

PF 12-NOV-1997; 97WO-JP04129.

PR 13-NOV-1996; 96JP-0302053.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dairiki K, Kuno K, Kume A, Sone T;

DR WPI; 1998-297617/26.

PT Peptides derived from Japanese cedar pollen antigens are
 PT immunotherapeutic agents - useful for allergy treatment and typing
 PT HLA class II molecules in allergy sufferers

PS Claim 12; Page 28; 50pp; Japanese.

XX This sequence represents residues 211-225 of the Cry j 1 protein, and
 CC is a peptide of the invention. The peptides are derived from Japanese
 CC cedar pollen antigens, and are used as immunotherapeutic agents in the
 CC treatment of allergy. The peptides can be used for identification and
 CC typing of the particular HLA class II molecules in an allergy sufferer,
 CC and also for peptide immunotherapy of an allergy. Using these peptides
 CC the immunotherapy can be targeted more specifically to the requirements
 CC of the individual patient, allowing more effective treatment of an
 CC allergy, including those patients for whom treatment with a conventional
 CC immunotherapeutic agent is ineffective.

SQ Sequence 15 AA;

Query Match 100.0%; Score 78; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.2e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15
 |||||
 DB 1 KSMKVTVAFNQFGPN 15

RESULT 3

AAM80349
 ID AAM80349 standard; peptide; 17 AA.

AC AAM80349;

DT 11-JAN-1999 (first entry)

DE Sugi allergen protein Cryj2 derived epitope for T cells.

KM T cell epitope; sugi allergen proteins Cryj1, Cryj2; treatment;
 KM sugi-pollenosis; allergic reaction; pollen.

OS Synthetic.

PN JP10259198-A.

PD 29-SEP-1998.

PF 22-DEC-1997; 97JP-0353448.

PR 24-DEC-1996; 96JP-0343441.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.

DR WPI; 1998-577037/49.

XX A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollenosis

PS Claim 7; Page 18; 21pp; Japanese.

CC AAM80339-58 represent epitopes for T cells, derived from the sugi
 CC allergen proteins Cryj1 (AAM80339-44, AAM80350-53 and AAM80356-58) and
 CC Cryj2 (AAM80345-49 and AAM80354-55). The peptides are useful for the
 CC treatment of sugi-pollenosis, an allergic reaction of the body to
 CC pollen.

SQ Sequence 17 AA;

Query Match 100.0%; Score 78; DB 19; Length 17;
 Best Local Similarity 100.0%; Pred. No. 4.2e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15
 |||||
 DB 2 KSMKVTVAFNQFGPN 16

RESULT 4

ID AAR45563 standard; Protein; 20 AA.

AC AAR45563;

DT 13-JUL-1994 (first entry)

DE Cry j 1 pollen allergen peptide CJI-22.

KM Japanese cedar; detection; allergy; treatment; diagnosis;
 KM T cell epitope; sensitivity.

OS Cryptomeria japonica.

PN WO9401560-A.

PD 20-JAN-1994.

PF 15-JAN-1993; 93WO-US00139.

PR 01-SEP-1992; 92US-0938990.
 PR 10-JUL-1992; 92WO-US05661.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J;
 XX WPI; 1994-035066/04.
 DR
 XX
 PT Antigens derived from Japanese cedar pollen allergen Cry j I -
 PT contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy
 XX
 PS Claim 1; Fig 13; 137bp; English.
 XX
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j I (amino acids 211-230). The peptide, CJI-22,
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen. It has enhanced therapeutic properties
 CC but reduced side effects compared to naturally occurring allergens.
 CC
 SQ Sequence 20 AA;
 SQ
 QY Query Match 100.0%; Score 78; DB 15; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5.8e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KSMKVTVAFNQFGPN 15
 DB 1 KSMKVTVAFNQFGPN 15
 RESULT 5
 ID AAR82512 standard; Protein; 20 AA.
 AC AAR82512;
 XX
 DT 15-APR-1996 (first entry)
 XX
 DE Cry j I Japanese Cedar pollen allergen peptide fragment (CJI-22).
 XX
 KM Cry j I; Japanese cedar pollen allergen; modified; drug production;
 KM allergy; Crytpomeria japonica.
 KM
 OS Crytpomeria japonica.
 XX
 PN WO9527786-A1.
 XX
 PD 19-OCT-1995.
 XX
 PF 06-APR-1995; 95WO-US04249.
 XX
 PR 06-DEC-1994; 94US-0350225.
 PR 08-APR-1994; 94US-0226248.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP;
 PI Shaked Z;
 PI
 DR WPI; 1995-366391/47.
 XX
 PT Modified Crytpomeria japonica (Cry j) I peptide(s) - useful for
 PT treating allergy to Japanese cedar pollen allergen or
 PT immunologically cross reactive allergens
 PT
 PS Disclosure; Figure 2; 60pp; English.
 XX
 CC Novel peptides of cry j I have been modified as a part of a
 CC preformulation scheme to develop an optimised drug product for
 CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j I,

CC modified and unmodified, are given in AAR82491-R82525. This peptide
 CC fragment corresponds to amino acids 211-230 of the allergen mature
 CC protein.
 CC
 SQ Sequence 20 AA;
 SQ
 QY Query Match 100.0%; Score 78; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5.8e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KSMKVTVAFNQFGPN 15
 DB 1 KSMKVTVAFNQFGPN 15
 RESULT 6
 ID AAW42144 standard; peptide; 20 AA.
 AC AAW42144;
 XX
 DT 16-JUN-1998 (first entry)
 XX
 DE T-cell epitope peptide 24 from Japanese cypress pollen antigen Chaol.
 XX
 KM Japanese cypress pollen; antigen; T-cell epitope; Chaol; Chaol2;
 KM diagnosis; allergy; spring tree pollen disease; pollinosis.
 KM
 OS Chamaeyparis obtusa.
 XX
 PN WO9747648-A1.
 XX
 PD 18-DEC-1997.
 XX
 PF 12-JUN-1997; 97WO-JP02031.
 XX
 PR 14-JUN-1996; 96JP-0153527.
 XX
 PA (MEIP) MEIJI MILK PROD CO LTD.
 XX
 PI Dairiki K, Kino K;
 PI
 DR WPI; 1998-052242/05.
 XX
 PT T-cell epitope peptide portion of Japanese cypress pollen antigens
 PT Chaol and Chaol2 - used for diagnosis and treatment of spring tree
 PT pollen disease
 PT
 PS Claim 1; Page 27-28; 71pp; Japanese.
 XX
 CC The present sequence represents a T-cell epitope peptide from Japanese
 CC cypress pollen antigen Chaol. The present invention describes peptides
 CC which correspond to the T-cell epitope sites on Japanese cypress pollen
 CC antigens Chaol and Chaol2. The peptides can be used as a reagent for the
 CC diagnosis of allergy to Japanese cypress pollen, and as an antigen in
 CC the treatment and prevention of spring tree pollen disease in which the
 CC pollinosis involves reactivity to Japanese cypress pollen.
 CC
 SQ Sequence 20 AA;
 SQ
 QY Query Match 100.0%; Score 78; DB 19; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5.8e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KSMKVTVAFNQFGPN 15
 DB 1 KSMKVTVAFNQFGPN 15
 RESULT 7
 ID AAR45591 standard; Protein; 26 AA.
 XX

AC AAR45591;
 XX 13-JUL-1994 (first entry)
 DT
 XX
 DE Cry j I pollen allergen peptide CJI-43.9.
 XX
 XX Japanese cedar; detection; allergy; treatment; diagnosis;
 KM T cell epitope; sensitivity.
 XX
 OS Cryptomeria japonica.
 XX
 PN WO9401560-A.
 XX
 PD 20-JAN-1994.
 XX
 PF 15-JAN-1993; 93WO-US00139.
 XX
 PR 01-SEP-1992; 92US-0938990.
 PR 10-JUL-1992; 92WO-US05661.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;
 DR WPI; 1994-035066/04.
 XX
 DT Antigen derived from Japanese cedar pollen allergen Cry j I -
 PT contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy
 XX
 PS Claim 76; Fig 18; 137pp; English.
 XX
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j I. The peptide, CJI-43.9, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 CC
 SQ Sequence 26 AA;
 Query Match 100.0%; Score 78; DB 15; Length 26;
 Best Local Similarity 100.0%; Pred. No. 7.8e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 KSMKVTVAFNQFGPN 15
 DB 1 KSMKVTVAFNQFGPN 15
 RESULT 8
 AAR45592
 ID AAR45592 standard; Protein; 26 AA.
 XX
 AC AAR45592;
 XX
 DT 13-JUL-1994 (first entry)
 XX
 DE Cry j I pollen allergen peptide CJI-43.10.
 XX
 KM Japanese cedar; detection; allergy; treatment; diagnosis;
 KM T cell epitope; sensitivity.
 XX
 OS Cryptomeria japonica.
 XX
 PN WO9401560-A.
 XX
 PD 20-JAN-1994.
 XX
 PF 15-JAN-1993; 93WO-US00139.
 XX
 PR 01-SEP-1992; 92US-0938990.
 PR 10-JUL-1992; 92WO-US05661.
 XX

PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;
 XX
 DR WPI; 1994-035066/04.
 XX
 PT Antigen derived from Japanese cedar pollen allergen Cry j I -
 PT contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy
 XX
 PS Claim 76; Fig 18; 137pp; English.
 XX
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j I. The peptide, CJI-43.10, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 CC
 SQ Sequence 26 AA;
 Query Match 100.0%; Score 78; DB 15; Length 26;
 Best Local Similarity 100.0%; Pred. No. 7.8e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 KSMKVTVAFNQFGPN 15
 DB 1 KSMKVTVAFNQFGPN 15
 RESULT 9
 AAR45594
 ID AAR45594 standard; Protein; 26 AA.
 XX
 AC AAR45594;
 XX
 DT 13-JUL-1994 (first entry)
 XX
 DE Cry j I pollen allergen peptide CJI-43.12.
 XX
 KM Japanese cedar; detection; allergy; treatment; diagnosis;
 KM T cell epitope; sensitivity.
 XX
 OS Cryptomeria japonica.
 XX
 PN WO9401560-A.
 XX
 PD 20-JAN-1994.
 XX
 PF 15-JAN-1993; 93WO-US00139.
 XX
 PR 01-SEP-1992; 92US-0938990.
 PR 10-JUL-1992; 92WO-US05661.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;
 XX
 DR WPI; 1994-035066/04.
 XX
 PT Antigen derived from Japanese cedar pollen allergen Cry j I -
 PT contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy
 XX
 PS Claim 76; Fig 18; 137pp; English.
 XX
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j I. The peptide, CJI-43.12, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 CC
 SQ Sequence 26 AA;

Query Match 100.0%; Score 78; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 7.8e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15
|||
DB 1 KSMKVTVAFNQFGPN 15

RESULT 10

AAR45589 standard; Protein; 28 AA.

AC AAR45589;

DT 13-JUL-1994 (first entry)

DE Cry j I pollen allergen peptide CUI-43.7.

KW Japanese cedar; detection; allergy; treatment; diagnosis;
PT T cell epitope; sensitivity.

OS Crytomeria japonica.

PN WO9401560-A.

PD 20-JAN-1994.

PF 15-JAN-1993; 93WO-US00139.

PR 01-SEP-1992; 92US-0938990.

PR 10-JUL-1992; 92WO-US05661.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;

DR WPI; 1994-035066/04.

PT Antigens derived from Japanese cedar pollen allergen Cry j I -
PT contain at least two T cell epitope(s), used to treat or diagnose
PT allergy

PS Claim 76; Fig 18; 137pp; English.

CC The sequence is that of an isolated peptide of the Japanese cedar
CC pollen allergen Cry j I. The peptide, CUI-43.7, can be used for
CC the treatment and diagnosis of allergies associated with Japanese
CC cedar pollen. It has enhanced therapeutic properties but reduced
CC side effects compared to naturally occurring allergens.

SQ Sequence 28 AA;

Query Match 100.0%; Score 78; DB 15; Length 28;

Best Local Similarity 100.0%; Pred. No. 8.5e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15
|||
DB 1 KSMKVTVAFNQFGPN 15

RESULT 11

AAR45590 standard; Protein; 28 AA.

AC AAR45590;

DT 13-JUL-1994 (first entry)

DE Cry j I pollen allergen peptide CUI-43.8.

KW Japanese cedar; detection; allergy; treatment; diagnosis;

KW T cell epitope; sensitivity.

OS Crytomeria japonica.

PN WO9401560-A.

PD 20-JAN-1994.

PF 15-JAN-1993; 93WO-US00139.

PR 01-SEP-1992; 92US-0938990.

PR 10-JUL-1992; 92WO-US05661.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;

DR WPI; 1994-035066/04.

PT Antigens derived from Japanese cedar pollen allergen Cry j I -
PT contain at least two T cell epitope(s), used to treat or diagnose
PT allergy

PS Claim 76; Fig 18; 137pp; English.

CC The sequence is that of an isolated peptide of the Japanese cedar
CC pollen allergen Cry j I. The peptide, CUI-43.8, can be used for
CC the treatment and diagnosis of allergies associated with Japanese
CC cedar pollen. It has enhanced therapeutic properties but reduced
CC side effects compared to naturally occurring allergens.

SQ Sequence 28 AA;

Query Match 100.0%; Score 78; DB 15; Length 28;

Best Local Similarity 100.0%; Pred. No. 8.5e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQFGPN 15
|||
DB 1 KSMKVTVAFNQFGPN 15

RESULT 12

AAR45593 standard; Protein; 28 AA.

AC AAR45593;

DT 13-JUL-1994 (first entry)

DE Cry j I pollen allergen peptide CUI-43.11.

KW Japanese cedar; detection; allergy; treatment; diagnosis;

PT T cell epitope; sensitivity.

OS Crytomeria japonica.

PN WO9401560-A.

PD 20-JAN-1994.

PF 15-JAN-1993; 93WO-US00139.

PR 01-SEP-1992; 92US-0938990.

PR 10-JUL-1992; 92WO-US05661.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;

DR WPI; 1994-035066/04.

PT Antigens derived from Japanese cedar pollen allergen Cry j I -

PT contain at least two T cell epitope(s), used to treat or diagnose
PT allergy

XX Claim 76; Fig 18; 137pp; English.

CC The sequence is that of an isolated peptide of the Japanese cedar
CC pollen allergen Cry j I. The peptide, CJI-43.11, can be used for
CC the treatment and diagnosis of allergies associated with Japanese
CC cedar pollen. It has enhanced therapeutic properties but reduced
CC side effects compared to naturally occurring allergens.

XX Sequence 28 AA;

SQ

Query Match 100.0%; Score 78; DB 15; Length 28;
Best Local Similarity 100.0%; Pred. No. 8.5e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15
DB 1 KSMKVTVAFNQFGPN 15

RESULT 13

AA45587
ID AAR45587 standard; Protein; 30 AA.

XX AAR45587;

AC AAR45587;

DT 13-JUL-1994 (first entry)

DE Cry j I pollen allergen peptide CJI-43.11.

XX Japanese cedar; detection; allergy; treatment; diagnosis;

KW T cell epitope; sensitivity.

XX Cryptomeria japonica.

OS WO9401560-A.

XX 20-JAN-1994.

PD 15-JAN-1993; 93WO-US00139.

XX 01-SEP-1992; 92US-0938990.

PR 10-JUL-1992; 92WO-US05661.

XX (IMMU-) IMMUNOLOGIC PHARM CORP.

PA Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;

PI WPI, 1994-035066/04.

XX Antigen derived from Japanese cedar pollen allergen Cry j I -

PT contain at least two T cell epitope(s), used to treat or diagnose

PT allergy

XX Claim 76; Fig 18; 137pp; English.

PS The sequence is that of an isolated peptide of the Japanese cedar

XX pollen allergen Cry j I. The peptide, CJI-43.11, can be used for

CC the treatment and diagnosis of allergies associated with Japanese

CC cedar pollen. It has enhanced therapeutic properties but reduced

CC side effects compared to naturally occurring allergens.

XX Sequence 30 AA;

SQ

Query Match

Best Local Similarity 100.0%; Score 78; DB 15; Length 30;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15

DB 1 KSMKVTVAFNQFGPN 15

RESULT 14

AA45588
ID AAR45588 standard; Protein; 30 AA.

XX AAR45588;

AC AAR45588;

DT 13-JUL-1994 (first entry)

DE Cry j I pollen allergen peptide CJI-43.6.

XX Japanese cedar; detection; allergy; treatment; diagnosis;

KW T cell epitope; sensitivity.

XX Cryptomeria japonica.

OS WO9401560-A.

XX 20-JAN-1994.

PD 15-JAN-1993; 93WO-US00139.

XX 01-SEP-1992; 92US-0938990.

PR 10-JUL-1992; 92WO-US05661.

XX (IMMU-) IMMUNOLOGIC PHARM CORP.

PA Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;

PI WPI, 1994-035066/04.

XX Antigen derived from Japanese cedar pollen allergen Cry j I -

PT contain at least two T cell epitope(s), used to treat or diagnose

PT allergy

XX Claim 76; Fig 18; 137pp; English.

OS The sequence is that of an isolated peptide of the Japanese cedar

CC pollen allergen Cry j I. The peptide, CJI-43.6, can be used for

CC the treatment and diagnosis of allergies associated with Japanese

CC cedar pollen. It has enhanced therapeutic properties but reduced

CC side effects compared to naturally occurring allergens.

XX Sequence 30 AA;

SQ

Query Match 100.0%; Score 78; DB 15; Length 30;

Best Local Similarity 100.0%; Pred. No. 9.2e-08;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSMKVTVAFNQFGPN 15

DB 1 KSMKVTVAFNQFGPN 15

RESULT 15

AA44687
ID AAW44687 standard; peptide; 30 AA.

XX AAW44687;

AC AAW44687;

DT 01-MAY-1998 (first entry)

DE T-cell epitope peptide #6 of sugi pollen antigen.

XX T-cell epitope; sugi pollen antigen; sugi pollinosis.

XX Synthetic.

OS Cryptomeria japonica.

XX JP10007700-A.

XX 13-JAN-1998.

XX 24-JUN-1996; 96JP-0163287.
 PF
 XX
 PR 24-JUN-1996; 96JP-0163287.
 XX
 XX
 PA (DAIL) DAICEL CHEM IND LTD.
 PA (MEIJ) MEIJI SEIKA KAISHA LTD.
 XX
 DR WPI; 1998-133630/13.
 XX
 PT T cell epitope peptide of sugi pollen antigen - useful in the
 PT treatment of sugi pollinosis
 XX
 PS Claim 1; Page 4; 14pp; Japanese.
 XX
 CC T-cell epitope peptides AAW4682-88 and their derivatives react with
 CC sugi pollinosis patient peripheral blood T lymphocytes. A composition
 CC prepared by combining at least 2 of the above peptides and/or their
 CC derivatives is used for the prevention and treatment of sugi
 CC pollinosis.
 XX
 SQ Sequence 30 AA;

Query Match 100.0%; Score 78; DB 19; Length 30;
 Best Local Similarity 100.0%; Pred. No. 9.2e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSMKVTVAFNQGFN 15
 |||||
 DB 11 KSMKVTVAFNQGFN 25

Search completed: April 20, 2003, 13:06:08
 Job time : 18.1974 secs

